ABSTRACT

Title of Dissertation: EFFECTS OF AGING ON MIDBRAIN AND CORTICAL SPEECH-IN-NOISE PROCESSING

Alessandro Presacco, Doctor of Philosophy, 2016

Dissertation directed by: Assistant Professor Samira Anderson,
Hearing and Speech Sciences Department

Professor Jonathan Z. Simon,
Department of Electrical and Computer Engineering,
Department of Biology and Institute for Systems Research

Older adults frequently report that they can hear what they have been told but cannot understand the meaning. This is particularly true in noisy conditions, where the additional challenge of suppressing irrelevant noise (i.e. a competing talker) adds another layer of difficulty to their speech understanding. Hearing aids improve speech perception in quiet, but their success in noisy environments has been modest, suggesting that peripheral hearing loss may not be the only factor in the older adult's perceptual difficulties. Recent animal studies have shown that auditory synapses and cells undergo significant age-related changes that could impact the integrity of temporal processing in the central auditory system. Psychoacoustic studies carried out in humans have also shown that hearing loss can explain the decline in older adults' performance in quiet compared to younger adults, but these psychoacoustic measurements are not accurate in describing auditory deficits in noisy conditions. These results would suggest that temporal auditory processing deficits could play an important role in explaining the reduced ability of older adults to process speech in noisy environments. The goals of this dissertation were to

understand how age affects neural auditory mechanisms and at which level in the auditory system these changes are particularly relevant for explaining speech-in-noise problems. Specifically, we used non-invasive neuroimaging techniques to tap into the midbrain and the cortex in order to analyze how auditory stimuli are processed in younger (our standard) and older adults. We will also attempt to investigate a possible interaction between processing carried out in the midbrain and cortex.

EFFECTS OF AGING ON MIDBRAIN AND CORTICAL SPEECH-IN-NOISE PROCESSING

By

Alessandro Presacco

Dissertation submitted to the Neuroscience and Cognitive Science program of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctorate of Philosophy
2016

Advisory Committee:

Assistant Professor Samira Anderson, Chair

Professor Jonathan Z. Simon

Professor Sandra Gordon-Salant

Associate Professor Matthew J. Goupell

DEDICATION

To my two beautiful daughters Angela Faith and Sofia Grace, whose births were the happiest moments of my life.

To my beautiful wife Kristen, who patiently waited for me to complete my Ph.D.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

Background

Introduction to neuroimaging techniques used to record Auditory Evoked Potentials

Event-related potentials (ERPs) are brain responses time-locked to events that can be represented by visual (e.g. visual flash), acoustical (e.g. clicks, tone bursts, etc.) or somatosensory (i.e. electrical stimulation of afferent peripheral nerve fibers) stimuli. Brain activity can be recorded using different invasive (e.g. single unit activity recorded with microelectrode arrays implanted into the cortex) and non-invasive (e.g. electroencephalographic (EEG) signals recorded from electrodes placed on the scalp) techniques. Invasive techniques have the obvious advantage of having an excellent spatial resolution which comes from the possibility of detecting the action potential (AP) of single neurons. A big disadvantage is related to the fact that surgery is needed to implant the microelectrodes (or grids of electrodes as in case of the Electrocorticogram (ECoG)) to record the electrical activity of the cells. This makes their application feasible mainly in animals or in humans that already have devices implanted in their brains to treat specific neurological disorders (i.e. epilepsy). Conversely, non-invasive techniques such as EEG and MEG (magnetoencephalography) have a very poor spatial resolution (> 100,000 neurons per sensor), but have exceptional temporal resolution, which can be estimated in the order of millisecond (ms). If spatial resolution is a critical component for a specific study, functional magnetic resonance imaging (fMRI) could be used, even though the temporal resolution drops down to a few seconds. The schematic illustrated in figure 1 (Waldert et al. 2009) shows how spatial and temporal resolutions vary with the invasiveness of the neuroimaging technique utilized.

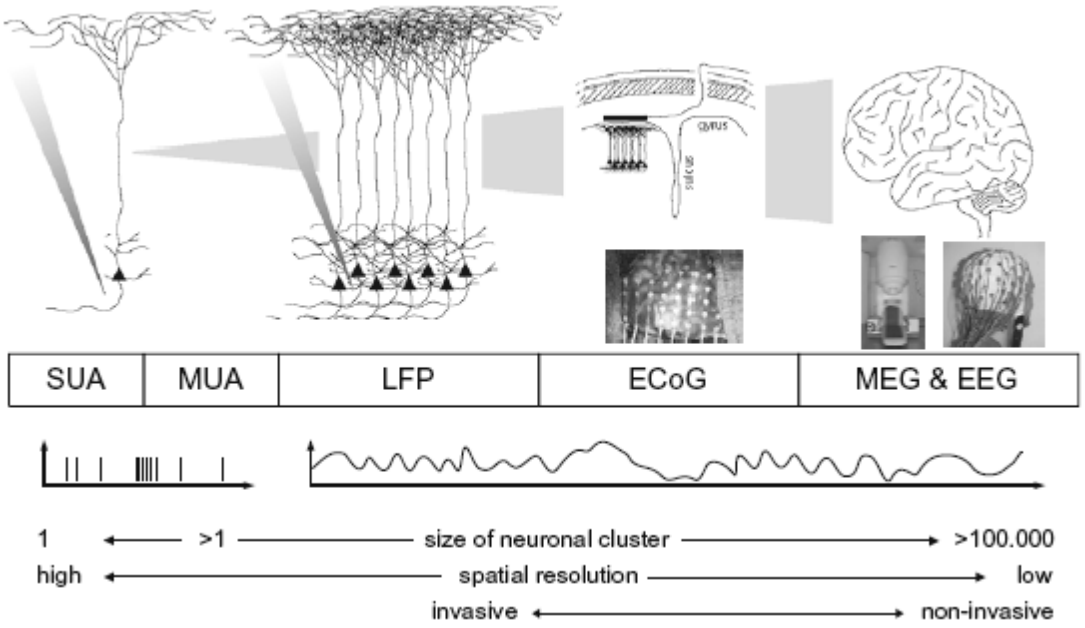

Fig. 1 Illustration of some of the most widely used neuroimaging techniques used to record evoked activity (Waldert et al. 2009). SUA = Single Unit Activity; MUA = Multi Unit Activity; LFP = Local Field Potential; ECoG = Electrocorticogram; MEG = Magnetoencephalography; EEG = Electroencephalography.

Basic principles of Electroencephalography (EEG)

EEG measures the voltage fluctuations (~μV) generated by the currents that flow during synaptic excitation of the dendrites of many pyramidal neurons in the cerebral cortex (Figure 2), which lie right under the skull and make up 80% of the brain's mass (Bear et al. 2006). This activity is recorded from electrodes (up to 256 for high density recordings) placed on the scalp. Given the very low electrical activity generated by each neuron and given the fact that the signal has to travel through different layers in the brain before being recorded by the electrodes, the voltage detected by each sensor is the result of the response of thousands of neurons that generate action potentials synchronously.

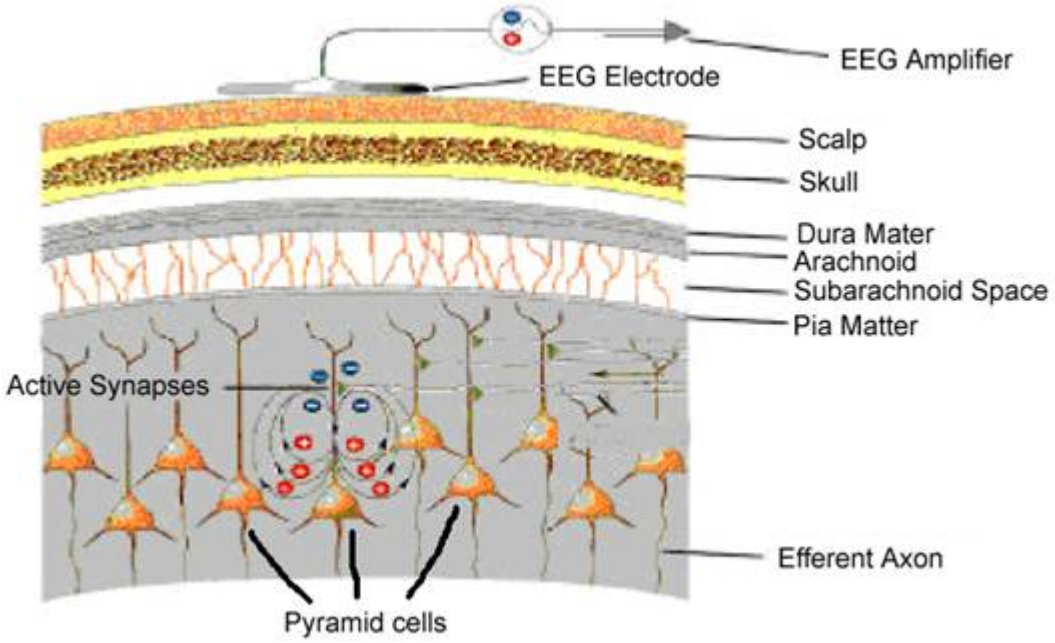

Fig. 2 Generation of very small electrical fields by synaptic currents in pyramidal cells (Bear et al. 2006).

Basic principles of Magnetoencephalography (MEG)

MEG measures the magnetic fields (~fT) that emanate from the head and are also generated by the electrical activity produced by neural activity (Figure 3). A clear advantage of MEG over EEG is the fact that magnetic fields are not volume conducted, which means that they pass undistorted through the scalp. Therefore, source localization would be feasible, even though overlapping of magnetic fields

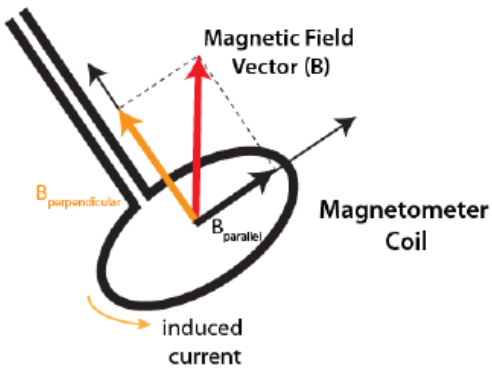

Fig. 3 Example of how a magnetic field passing through a conductive coil induces current that is directly proportional to the perpendicular component of the field (Cheyne and Papanicolaou 2006).

(superposition) is still possible, making the interpretation of topographical patterns ambiguous. Given the extremely low strength of the magnetic field generated by neurons (10^{-9} times lower than the Earth field), Superconductive Quantum Interference Devices (SQUID) are used. These detectors are submersed into liquid Helium in order to keep their temperature down to 4.2˚K (-269˚ C), which leads to superconductivity (see Figure 4). In order to increase the signal-to-noise ratio (SNR), these detectors are often designed as gradiometers, which consist of two coils wound in opposite directions separated by a small distance. The most commonly used gradiometer is the first order, which reduces the noise by 99%, while brain activity can be reduced by as much as 50%. MEG machines are usually placed in magnetically shielded rooms to reduce the environmental noise, which is much stronger than the magnetic fields generated by brain activity. Reference sensors are also used to further reduce external noise.

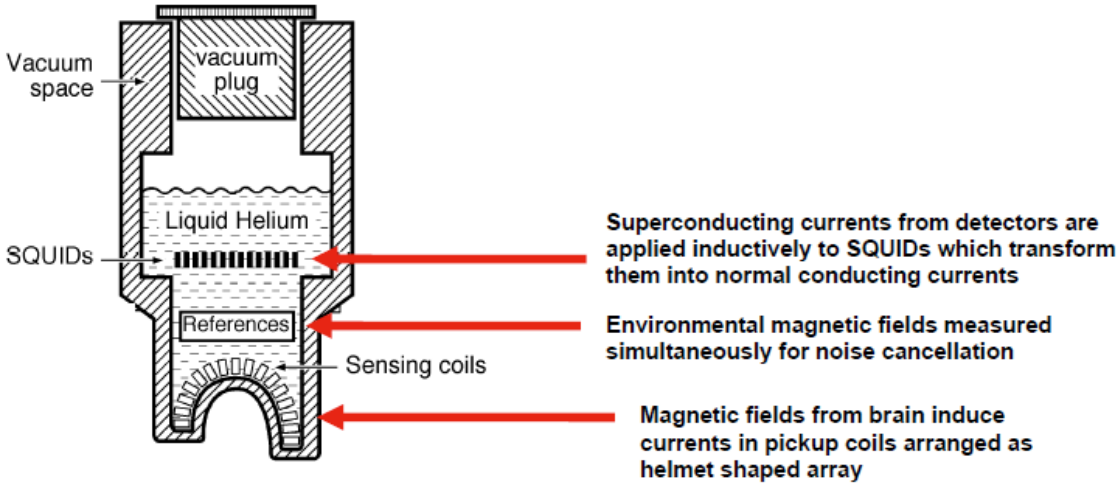

Fig. 4 Schematic of the QUIDS detectors submersed in liquid Helium. Reference sensors are used to reduce environmental noise (Cheyne and Papanicolaou 2006).

Classification of Auditory Evoked Potentials

Auditory evoked potentials (AEPs) are a subclass of ERPs and are evoked by using a variety of auditory stimuli (i.e. clicks, tone bursts). The AEPs that are recorded from the top of the head originate from different structures within the brain (i.e. the auditory cortex, inferior colliculus, the auditory cranial nerve VIII). AEPs are divided into 3 major groups, based on their latency:

1) ABR (Auditory Brainstem Response)
2) MLR (Middle Latency Response)
3) LLR (Late Latency Response)

Figure 5 (Michelini et al. 1982) shows the 3 major groups of AEPs. "P" stands for positive peak and "N" for negative peak.

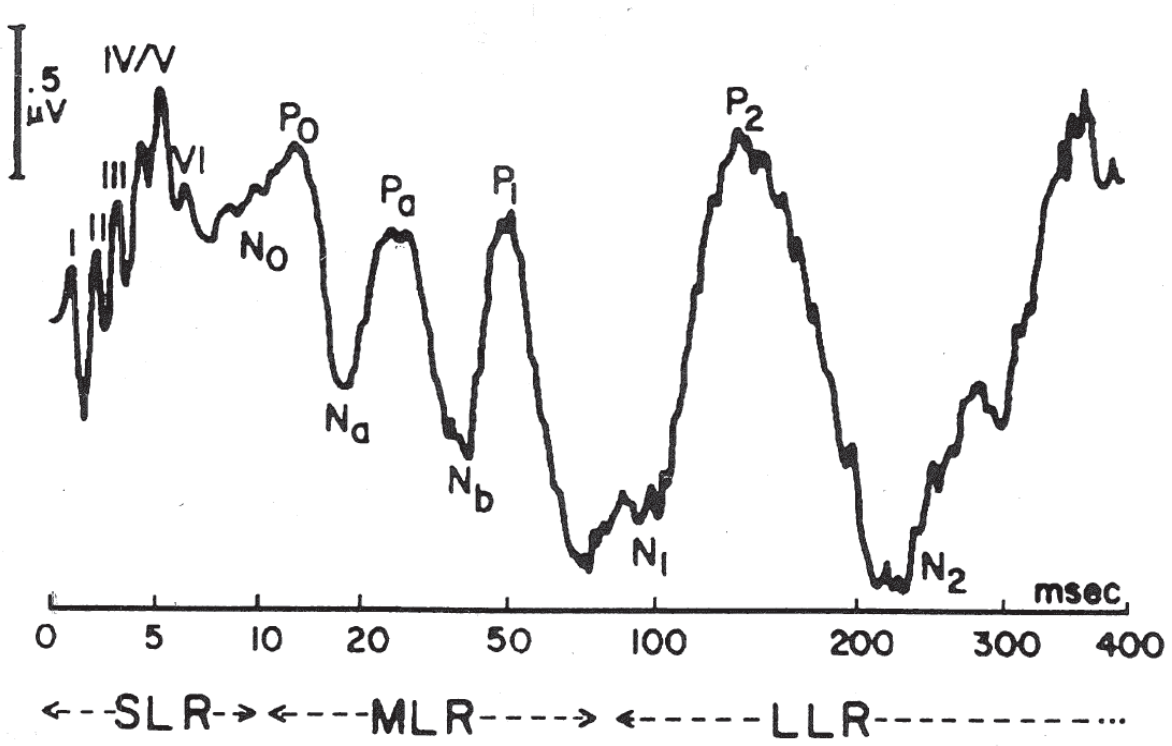

Fig. 5 Transient auditory evoked potential elicited by a single acoustic click stimulus (SLR is equivalent to ABR). The 3 different groups are divided based on their latency.

Auditory Brainstem Response (ABR)

ABRs occur in the first 10 ms from the stimulus onset. They are characterized by 5 major peaks, which are labelled based on their latency:

1. Peak I, which occurs around 1 ms post stimulus onset and originates in the cranial nerve VIII
2. Peak II, which occurs around 2 ms post stimulus onset
3. Peak III, which occurs around 3 ms post stimulus onset
4. Peak IV, which occurs around 4 ms post stimulus onset
5. Peak V, which occurs around 5 ms post stimulus onset

These peaks are believed to be mainly exogenous and are minimally affected by the mental status of the subject. They are usually elicited with stimuli with a rate of 10-20 per second, as higher rates tend to suppress ABR responses. The filter settings are generally set at approximately 100 - 3000 Hz and a minimum of 1000 sweeps are usually necessary to have an acceptable signal-to-noise ratio.

Possible clinical applications are in the screening and assessment of infants and young children with hearing loss, as well as in the neurophysiologic assessment of pediatric and adult patients who may have lesions associated with their auditory nerve VIII.

Auditory Middle Latency Response (MLR)

MLRs occur between 10 and 80 ms from the stimulus onset. They are mainly characterized by 3 components:

1. Na (negative peak following ABR peak V), which mainly originates in upper brainstem.
2. Pa (positive peak at about 30 ms), which originates in the auditory cortex bilaterally.
3. P1 (also called Pb or P50), which originates in the auditory cortex.

MLRs are believed to be partially affected by the mental condition of the subject. The filter settings are generally set at 10 - 1500 Hz range and a minimum number of 500 sweeps is usually necessary to have an acceptable SNR.

Possible clinical applications include: anesthesia, sleep monitoring, neuro-psychiatric testing, auditory development, evaluation of hearing impairment and its rehabilitation.

Auditory Late Latency Response (LLR)

LLRs occur after 80 ms from the stimulus onset. They are characterized by the following positive and negative peaks:

1. N1, which occurs at around 100 ms from stimulus onset and originates in the auditory cortex (Heschl's gyrus).
2. P2, which occurs at around 200 ms from stimulus onset and originates in the auditory cortex (Heschl's gyrus).

3. P300, which occurs around 300 ms. This ERP can divided into two compoenents: P3a (localized within frontal cortex and anterior cingulate gyrus) and P3b (localized in the superior parietal lobule and the posterior part of the cingulate gyrus) (Wronka et al. 2012). These large positive peaks occur only when a subject is actively engaged in the detection of a target stimulus, such as in an oddball paradigm, where deviant tones occur with a small probability rate (~20%).

The analog input filter settings are generally set at 1-30 Hz and due to the higher SNR, 200 sweeps (or less) are usually recorded. Possible clinical applications include evaluation of schizophrenia, autism, depression, dementia, cognitive disorders and encephalopathy.

Frequency Following Response (FFR)

A particular type of AEP is the auditory steady-state evoked response (ASSR), which is a sustained activity elicited in response to high-stimulation rate stimuli. This high stimulation rate causes the auditory responses to overlap, generating a convolved response that will have characteristics similar to the stimulation rate. The 40 Hz convolved response is one of the most widely studied ASSRs. It was first recorded by Galambos et al. (1981) who used click trains at a mean rate of 40 Hz (25 ms apart) to generate a convolved response that showed a resonance at 40 Hz. Figure 6 from Galambos et al. (1981) shows how this fast stimulation rate elicits an evoked response every 25 ms; these responses eventually overlap giving rise to a 40 Hz convolved response. Frequency following responses (FFRs) are a subcategory of AASRs and can be defined as brainstem responses to sustained speech and speech-like stimuli (Aiken and Picton 2008) whose putative

generator is believed to reside in the inferior colliculus (Galbraith et al. 2000; Smith et al. 1975). FFRs are usually recorded using stimuli with alternating polarity in order to reduce cochlear microphonic and electrical stimulus artifact (Aiken and Picton 2008). The responses to the two polarities are averaged to extract the envelope of the convolved response, which will have a period corresponding to the fundamental frequency of the auditory stimulus used to evoke the auditory response. FFRs can be used to investigate temporal processing deficits at the level of the midbrain.

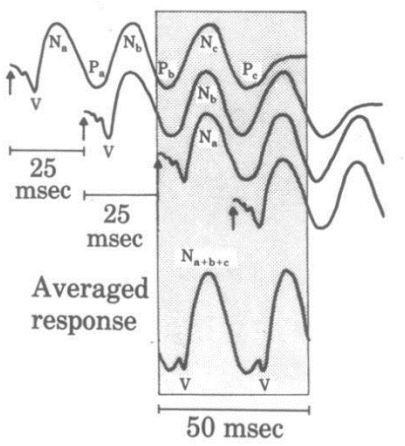

Fig. 6 Example of 40 Hz auditory activity first recorded by Galambos et al. (1981). The 50 ms convolved response shown at the bottom of the figure has a period of 25 ms, which perfectly "follows" the stimulation rate of the train of clicks.

Evoked activity related to the cocktail party problem

One of the most difficult tasks that the brain routinely performs with apparent little effort is the segregation of physical sounds that are linearly mixed together. The process by which the auditory system separates the individual sounds in the environment is called auditory scene analysis (ASA) (Bregman 1990). This is a mathematically ill-posed problem, as the sounds that enter the ear linearly mix and there are potentially several sets of sounds that occur in the real world, that could be potentially consistent with the mixture of input sounds (McDermott 2009). Related to ASA is the cocktail party problem (McDermott 2009), in which an individual faces the challenge of extracting the target stimulus (foreground) from a complex mixture of irrelevant stimuli (background). Attention plays a critical role in this task, because in order to bind together all the relevant features of a speech stream, at least

one feature of the target stream must be recognized by the listener in order to have access to the other temporal coherent features of the target speech (Shamma et al. 2011). It follows that if a listener is instructed to attend to a specific talker while trying to ignore the background noise, the evoked activity elicited in the brain should have some critical characteristic present in the target speech. Recent MEG studies (Ding and Simon 2013; 2012) showed that the brain phase locks to the envelope of the speech to which the individual is instructed to attend. Figure 7 shows a simple graphical representation of the task that Ding and Simon (2013; 2012) used in their experiment: selective attention determines to which speech envelope the brain will phase-lock.

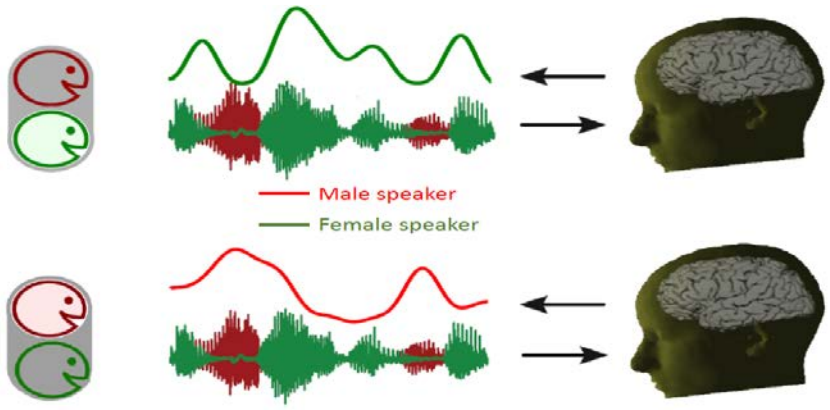

Fig. 7 Graphical representation of the auditory task. Subjects were instructed to attend to either the male speaker (red) or to the female speaker (green), while trying to ignore the competing talker. The brain response was used to reconstruct the envelope of the speech stimulus to which the participant was instructed to attend.

Reconstructing the speech envelope: the backward model

Ding and Simon (2013; 2012) reconstructed the speech envelope by using the backward model approach (Figure 8).

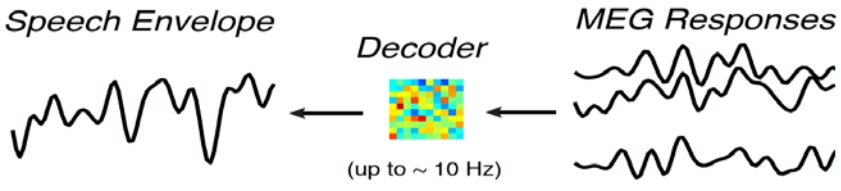

Fig. 8 Representation of the backward model used to reconstruct the speech envelope from MEG response.

MEG responses recorded from "*N*" channels were fed into a blind source separation algorithm that extracted up to "*N*" independent components, which were ordered based on the percentage of variability of the data that they can explain. For instance, the first component would be the one that explains most of the variability of the neural response. If an auditory stimulus is used, we would expect the first component to have the strongest representation of the auditory neural response, while as we look at lower ranked components, we should expect to see less contribution from the auditory cortices. Figure 9 shows the field distribution associated with the first component extracted from one subject, where the presence of two dipoles (one for each auditory cortex) is evident.

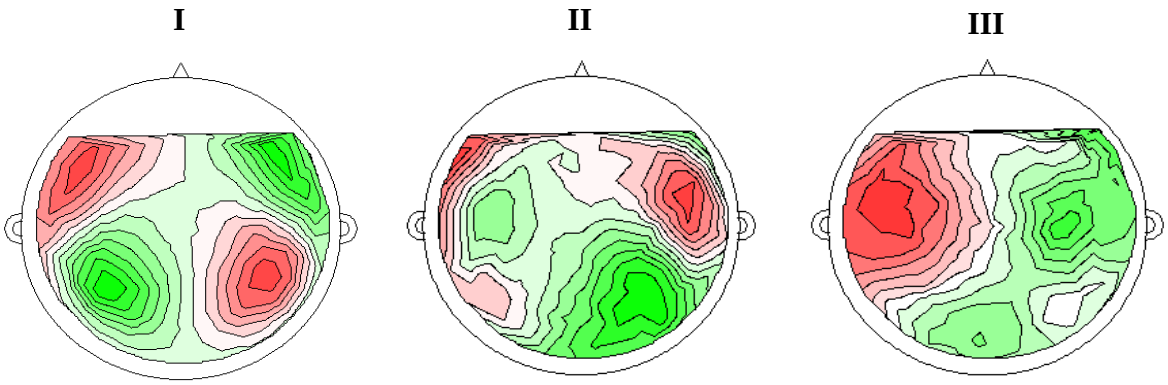

Fig. 9 Field distribution (fT) of the first three components. Note how two dipoles are well represented in both hemispheres at the level of the auditory cortex in the first component, whereas the second and third components might represent the sum or the difference between independent sources (e.g. M50 and M100).

Reverse correlation: the forward model

The forward problem is applied when acoustic stimuli are used to predict the evoked responses elicited by the sound. This technique can be used to estimate the spatio-temporal response field (or receptive field if we talk about single neural activity) (STRF), which is a mathematical kernel function. STRF is basically a linear model in the spectral-temporal domain that best describes the linear mapping between the stimulus spectrogram and the observed neural response (David et al. 2009). The importance of the forward model is due to the fact that only the STRF can reveal the timing and spatial information of the

neural encoding process (Ding and Simon 2012). This is because the STRF efficiently represents the way that the spectrotemporal features of speech are transformed into cortical responses. As described by Ding and Simon (2012) in their study (Figure 10), the STRF represents, at any frequency considered in the analysis during the time window investigated, the neural response to the features carried in the acoustic stimulus.

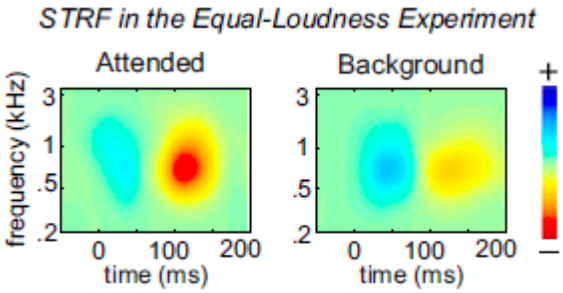

Fig. 10 STRF (arbitrary scale) of the attended and unattended speech reported by Ding and Simon (2012). Note how the two spectrograms differ from each other at around 100 ms, indicating a higher representation for the attended stimulus.

Introduction to proposed experiments

The ability to track and understand speech in the presence of interfering speakers is one of the most complex communication challenges experienced by humans. In a complex auditory scene, both humans and animals show a remarkable ability to detect and recognize individual auditory objects, a necessary step in the process of stream segregation. One key question is how the brain is able to transform the noise-corrupted signal into a reliable neural representation suitable for speech recognition. Recent results using magnetoencephalography (MEG) (Ding and Simon, 2012a; 2013; Ding et al. 2014) suggest that this representation may occur in the auditory cortex (AC) through adaptive neural encoding. Specifically, they demonstrated that low frequency auditory cortical activity is reliably synchronized to the slow temporal modulations of speech, even when the

background noise is significantly stronger than the speech signal. It has been suggested that the precision of this temporal synchronization is critical for achieving good ability to recognize speech in noise.

The abovementioned studies were carried out in young adults, but this temporal synchronization has not yet been demonstrated in older adults, who represent a critical age group as it is estimated that the number of adults 60 and older will dramatically increase in the next 10 years, bringing about a greater prevalence of communication problems (Lin et al. 2011). Older adults frequently report they can hear what is said but cannot understand the meaning, especially in noise. These communication difficulties have a significant social impact on older adults, as several studies have shown strong correlations among hearing loss and depression (Carabellese et al. 1993; Herbst and Humphrey 1980; Kay et al. 1964; Laforge et al. 1992) and cognitive impairment (Gates et al. 1996; Lin et al. 2013; Uhlmann et al. 1989). Although amplification benefits those who wear hearing aids (Humes et al. 2001), it may not improve speech understanding in noise for older adults, as increased audibility does not restore temporal precision degraded by aging (Tremblay et al. 2003).

Evidence of age-related deficits in auditory temporal processing has been found in psychoacoustic studies (Fitzgibbons and Gordon-Salant 1996; Frisina and Frisina 1997; Gordon-Salant et al. 2006; He et al. 2008; Pichora-Fuller and Schneider 1991; Schneider and Hamstra 1999), at the single neuron level from various nuclei of the auditory pathway in animal models (Recanzone et al. 2011; Schatteman et al. 2008; Walton et al. 1998), and in electrophysiological studies in humans and animals (Anderson et al. 2012; Clinard and Tremblay 2013; Lister et al. 2011; Parthasarathy and Bartlett 2011). A number of rodent studies support the theory that this degradation of temporal precision may be attributed to

a significant decrease of inhibitory functions and a consequent loss of balance between excitatory and inhibitory processes in the dorsal cochlear nuclei (Caspary et al. 2005; Schatteman et al. 2008; Wang et al. 2009), inferior colliculi (IC) (Caspary et al. 1995), spiral ganglion neurons (Tang et al. 2014), and auditory cortices (de Villers-Sidani et al. 2010; Hughes et al. 2010; Juarez-Salinas et al. 2010; Overton and Recanzone 2016).

This dissertation will evaluate the effects of aging on temporal synchronization of speech in quiet at the midbrain level (First study) and in the presence of competing speech (Second and third study) in the cortex and in the midbrain using neuroimaging techniques such as EEG and MEG. The effect of noise on speech was evaluated using meaningful (competing English talker) and meaningless (competing Dutch speaker) competing noise, as older adults make use of the type of noise to compensate for their problems in understanding speech in noise (Tun et al., 2002). We will also evaluate the relationships between MEG and EEG measures of temporal synchronization. This is particularly important in light of a recent study that has shown a compensatory central gain increases strong enough to help restore the representation of the auditory object at the cortical level, even in absence of a brainstem response (Chambers et al. 2016). Our central hypothesis is that the neural encoding of the speech envelope in competing noise is degraded in older adults compared to younger adults, even when the older adults have “normal hearing” based on commonly used clinical assessments. Thus, in the presence of noise, accurate cortical temporal encoding becomes a prerequisite for speech recognition, and listeners who fare better at extracting basic acoustic modulations also perform better on measures of speech-in-noise recognition. We hypothesize that faulty encoding in midbrain can affect the segregation of the speech streams at the cortical level. This study may have a significant

impact on the treatment of hearing difficulties in older adults, as the implications of the results would suggest that strategies extending beyond enhancing audibility should be adopted. For example, given that training improves neural synchrony in rats, with a partial reversal of behavioral deficits (de Villers-Sidani et al., 2010), the identification of neural processing deficits in aging humans may guide future treatment interventions.

CHAPTER II

First study

Objectives

We investigated aging effects on the envelope of the frequency following response to dynamic and static components of speech. Older adults frequently experience problems understanding speech, despite having clinically normal hearing. Improving audibility with hearing aids provides variable benefit, as amplification cannot restore the temporal precision degraded by aging. Previous studies have demonstrated an age-related delay in subcortical timing specific to the dynamic, transition region of the stimulus. However, it is not known whether this delay is mainly due to a failure to encode rapid changes in the formant because of central temporal processing deficits or as a result of cochlear damage that reduces audibility for the high-frequency components of the speech syllable. To investigate the nature of this delay, we compared subcortical responses in younger and older adults with normal hearing to the speech syllables /da/ and /a/, hypothesizing that the delays in peak timing observed in older adults are mainly caused by temporal processing deficits in the central auditory system.

Introduction

The number of adults 60 years and older will dramatically increase in the next 10 years, bringing about a greater prevalence of communication problems (Lin et al. 2011). Older adults frequently report they can hear what is said but cannot understand the meaning, especially in noise. These communication difficulties have a significant social

impact on older adults, as several studies have shown strong correlations among hearing loss and depression (Carabellese et al. 1993; Herbst and Humphrey 1980; Kay et al. 1964; Laforge et al. 1992) and cognitive impairment (Gates et al. 1996; Lin et al. 2013; Uhlmann et al. 1989). Although amplification benefits hearing aid users (Humes et al. 2001), it may not improve speech understanding in noise for older adults, as increased audibility does not restore temporal precision degraded by aging (Tremblay et al. 2003). Evidence of age-related deficits in auditory temporal processing has been found in psychoacoustics studies (Fitzgibbons and Gordon-Salant 1996; Frisina and Frisina 1997; Gordon-Salant et al. 2006; He et al. 2008; Pichora-Fuller and Schneider 1991; Schneider and Hamstra 1999), at the single neuron level from various nuclei of the auditory pathway in animal models (Recanzone et al. 2011; Schatteman et al. 2008; Walton et al. 1998) and in electrophysiological studies in humans and animals (Anderson et al. 2012; Clinard and Cotter 2015; Clinard and Tremblay 2013; Lister et al. 2011; Parthasarathy and Bartlett 2011). A number of rodent studies support the theory that this degradation of temporal precision may be attributed to a significant decrease of inhibitory functions and a consequent loss of balance between excitatory and inhibitory processes in the dorsal cochlear nuclei (Caspary et al. 2005; Schatteman et al. 2008; Wang et al. 2009), inferior colliculi (IC) (Caspary et al. 1995), spiral ganglion neurons (Tang et al. 2014), and auditory cortices (de Villers-Sidani et al. 2010; Hughes et al. 2010; Juarez-Salinas et al. 2010). Anderson et al. (2012) found peak latency delays in older adults' responses evoked by a consonant–vowel (CV) syllable, but only in responses to the formant transition region of the syllable and not in responses to the steady state. What remains to be clarified is if this delay arises from central temporal processing deficits that reduce the precision of encoding

the rapidly changing formants in the CV transition or from an onset delay originating in the cochlea. The stop consonant burst in the /da/ syllable used in previous studies contains high-frequency energy (Anderson et al. 2012; Clinard and Tremblay 2013; Vander Werff and Burns 2011). Because older adults typically have higher audiometric thresholds than younger adults in the high frequencies, even when their hearing would be classified as clinically normal, the delay in latency may be due to reduced audibility of the /da/ consonant. We hypothesized that impaired auditory temporal processing is the main cause of the delayed neural timing for the rapidly changing formants in the dynamic regions of speech in older adults. To test our hypothesis, we recorded frequency following responses (FFRs) in younger and older adults to the speech syllables /da/ and /a/. We expected to find latency delays for the CV transition of the /da/, but not for the same time region of the /a/, in older adults compared with younger adults. If decreased audibility is a factor in the delayed latencies for the /da/, however, we would expect that latencies for the /da/ would be earlier than those for the /a/ in the younger group given cochlear tonotopicity, but in the older group, damage in the basal end of the cochlea would lead to similar latencies for the /da/ and /a/.

MATERIALS AND METHODS

Participants

Participants comprised 15 younger adults (21 to 30 years old; mean ± standard error [SE], 24.4 ± 0.65 years; two males) and 15 older adults (60 to 68 years old; mean ± SE, 63.73 ± 0.66 years; four males) recruited from the Maryland, Washington, DC, and Virginia areas. All participants had clinically normal hearing defined as follows: (1) air conduction

thresholds ≤ 25 dB HL from 125 to 4000 Hz bilaterally and (2) no interaural asymmetry > 15 dB HL at two or more adjacent frequencies. In addition, all subjects had normal click-evoked brainstem response latencies (wave V < 6.8 msec; Otto and McCandless (1982)) measured using a 100-μs click stimulus presented at 80 dB peSPL at a rate of 31.3 Hz. No participants reported a history of neurological disorders, and all were native speakers of English. Because of the established effects of musicianship on subcortical auditory processing (Bidelman and Krishnan 2010; Parbery-Clark et al. 2012), professional musicians were excluded and the groups were matched on self-rated instrument proficiency [$t(28) = 0.874$, $p = 0.389$]. The two groups had normal IQs (≥ 85 on the Wechsler Abbreviated Scale of Intelligence; Zhu and Garcia (1999)) and were matched on IQ [$t(28) = 1.691$, $p = 0.102$] and sex (Fisher exact, $p = 0.651$). In addition, the older adults were screened for dementia using the Montreal Cognitive Assessment (Nasreddine et al. 2005). See Figure 1 for average audiometric thresholds and Table 1 for mean and SEs of the test scores, audiometric thresholds, and click-evoked response latencies. All procedures were reviewed and approved by the Institutional Review Board of the University of Maryland. Participants gave informed consent and were paid for their time.

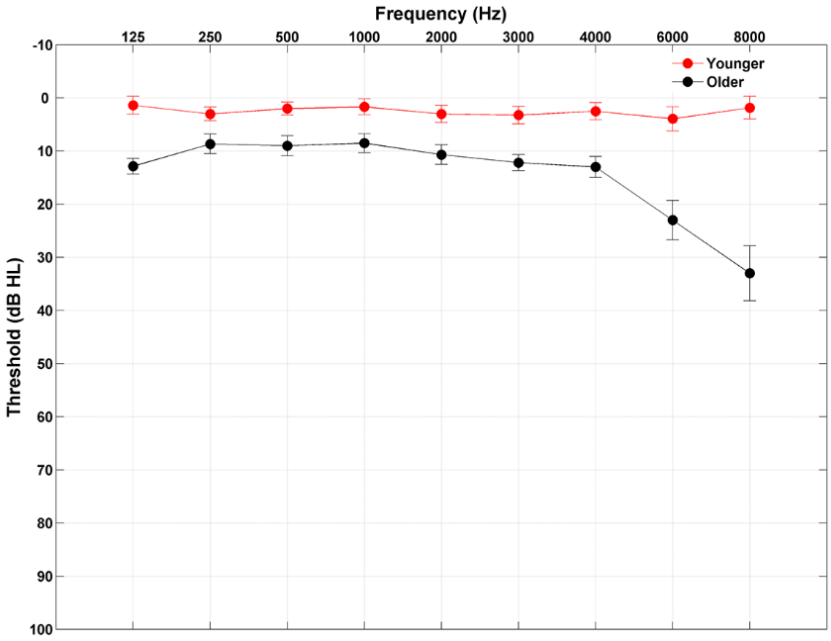

Fig. 11 Audiogram (mean ± 1SE) of the grand averages of younger and older adults.

Table 1 Group means (with standard errors) for the younger and older adults for pure tone averages (0.125–4 kHz), age, click latency wave I and wave V, Wechsler Abbreviated Scale of Intelligence IQ, and Montreal Cognitive Assessment scores

	Pure Tone Average for 0.125–8 kHz (dB HL)	Age (yr)	Click Latency Wave I (msec)	Click Latency Wave V (msec)	Wechsler Abbreviated Scale of Intelligence IQ (Standard Score)	Montreal Cognitive Assessment (Standard Score)
Younger (n = 15)	2.519 (0.390)	24.40 (0.65)	1.69 (0.03)	5.65 (0.05)	110.80 (3.14)	Not applicable
Older (n = 15)	14.552 (0.774)	63.73 (0.66)	1.73 (0.06)	6.04 (0.07)	119.13 (3.79)	25.86 (0.44)
Older (n = 6)	10.093 (0.702)	63.33 (1.22)	1.67 (0.09)	5.95 (0.10)	114.33 (7.07)	25.83 (0.40)

Electrophysiology

Stimuli. Two speech syllables were used. (1) A 170-ms /da/ was synthesized at a 20 kHz sampling rate with a Klatt-based synthesizer (Klatt 1980). After an initial 10-ms stop burst in the syllable, voicing remained constant with a fundamental frequency (F_0) of 100 Hz. During the 50 ms transition from the /d/ to the /a/, the lower three formants shifted (F_1, 400

→ 720 Hz; F_2, 1700 → 1240 Hz; F_3, 2580 → 2500 Hz) but stabilized for the 120-ms steady-state vowel portion. The fourth through sixth formants (F_4 through F_6) remained constant over 170 ms at 3300, 3750, and 4900 Hz, respectively. A waveform of the /da/ is presented in Figure 12 along with average responses in younger and older adults. (2) A 170-ms /a/ was synthesized at a 20 kHz sampling rate with a Klatt-based synthesizer (Klatt 1980). This syllable had a fundamental frequency (F_0) of 100 Hz and, in contrast to the /da/, all the first six formants stayed constant throughout the whole presentation of the stimulus (F_1 = 720 Hz; F_2 = 1240 Hz; F_3 = 2500 Hz; F_4 = 3300 Hz; F_5 = 3750 Hz; F_6 = 4900 Hz). A waveform of the /a/ is presented in Figure 13 along with average responses in younger and older adults. Despite the absence of a transition region in the /a/, the peaks occurring between 22 and 62 ms are marked as transition in order to compare this time window with the equivalent transition segment observed in the /da/. The /da/ was chosen because it includes a formant-changing transition and a steady-state region, and the /a/ was chosen because the formants are constant throughout the waveform. Both speech syllables were presented binaurally using Intelligent Hearing Systems SmartEP system (IHS; Miami, FL) with alternating polarities at 80 dB SPL at a rate of 4.3 Hz through electromagnetically shielded insert earphones (ER-3; Etymotic Research) at a sampling rate of 13.3 kHz.

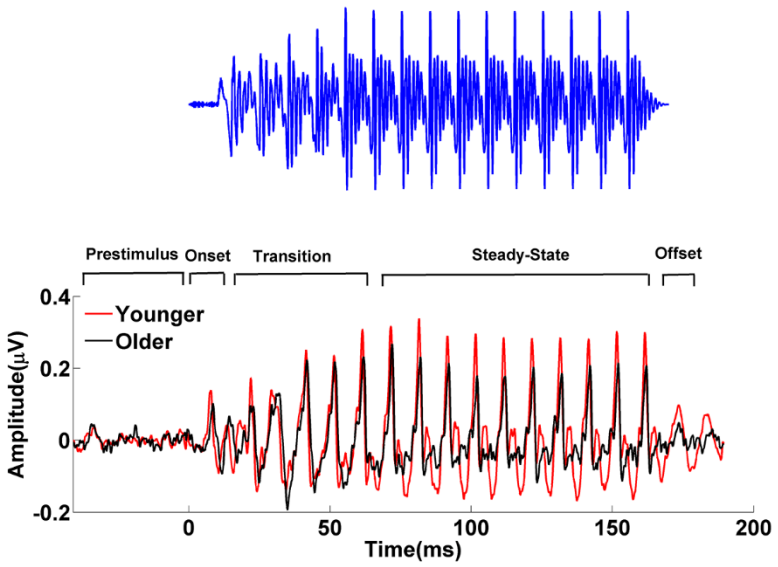

Fig. 12 Stimulus waveform (top) and average brainstem responses to /da/ in the younger (red) and older (black) adults (bottom). The prestimulus and response regions are labeled with respect to the onset, formant transition, and steady-state vowel of the stimulus.

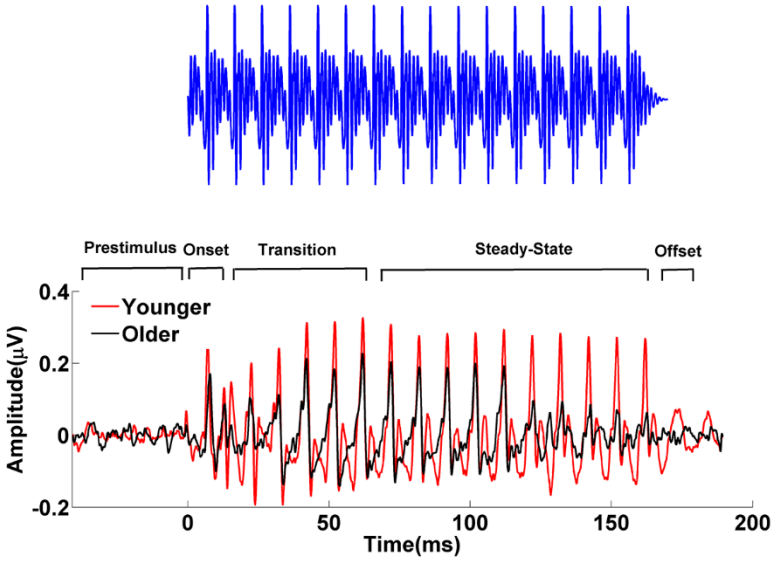

Fig. 13 Stimulus waveform (top) and average brainstem responses to /a/ in the younger (red) and older (black) adults (bottom). The steady-state region in the /a/ response starts at ~20 ms, but for comparison purposes with /da/, the transition region, in addition to the prestimulus, onset, steady-state, and offset regions, is also marked on this figure.

Recording. Subcortical responses were recorded in an electrically-shielded double-walled sound-attenuating chamber. The online filter was set at 50 – 3000 Hz. A vertical montage of four Ag-AgCl electrodes (Cz, active, forehead ground, average earlobes as reference) was used. During the recording session (~1hr), participants sat in a recliner and watched a silent, captioned movie of their choice to facilitate a relaxed yet wakeful state. The order of syllable presentation was randomized. Six thousand artifact-free sweeps were recorded for each speech syllable from each participant.

Data analysis. Sweeps with amplitudes in the ±31 μV range were retained and averaged in real time and then processed off-line using MATLAB (Mathworks, version R2011b). Responses were digitally bandpass-filtered offline from 70 to 2000 Hz using a Butterworth filter in order to minimize the effects of low frequency oscillations that originate from the cortex, given the subcortical origin of the FFR (Galbraith et al. 2000; Smith et al. 1975). Although some neurons in the auditory cortex are capable of phase locking up to 100 Hz (Wallace et al. 2000), the number of these neurons is much smaller than can be found in subcortical structures (Bartlett et al. 2007; Wang et al. 2008), and therefore the cortical contribution to the response is likely to be minimal. The time window for each sweep was -41 to 189 ms referenced to the stimulus onset. The final average was computed by averaging the 6000 sweeps (3000 of each polarity). One final average response was created for subsequent analysis, for which the two polarities were added to minimize the influence of cochlear microphonic and stimulus artifact on the response. Adding alternating polarities also minimizes the temporal fine structure while maximizing the envelope response (Aiken and Picton 2008; Campbell et al. 2012; Chimento and Schreiner 1990). Adding alternating polarities to extract the envelope introduces rectification-related

distortion at the level of the cochlea into the response which may produce energy at integer multiples of the F_0. Previous studies have shown that energy at the F_0 makes the strongest contribution to the envelope of the FFR (Aiken and Picton 2006; Greenberg et al. 1987). Added responses were used to analyze latency, amplitude, frequency and time-frequency representation of the F_0 and lower harmonics, and phase-locking value.

Analyses

Timing. To analyze aging effects on neural timing, we extracted peaks in the subcortical responses in two steps. First, a function written in MATLAB identified the peaks that were closest to the expected latency. These latencies were chosen based on the peaks extracted from the group average and from latencies obtained in previous studies (Anderson et al. 2012; Parbery-Clark et al. 2012). Then a trained peak picker, blind to participant group, confirmed each peak identification and made changes where appropriate. This identification provided the latency and amplitude of each peak. Peaks were labeled according to a reference latency of the speech syllable /a/ (i.e., a peak occurring 32–33 ms after onset would be called "peak 32"; see Fig. 14). The onset peak was identified as peak 8, transition peaks were 32, 42, and 52, and steady-state peaks were 62, 72, 82, 92, 102, and 112.

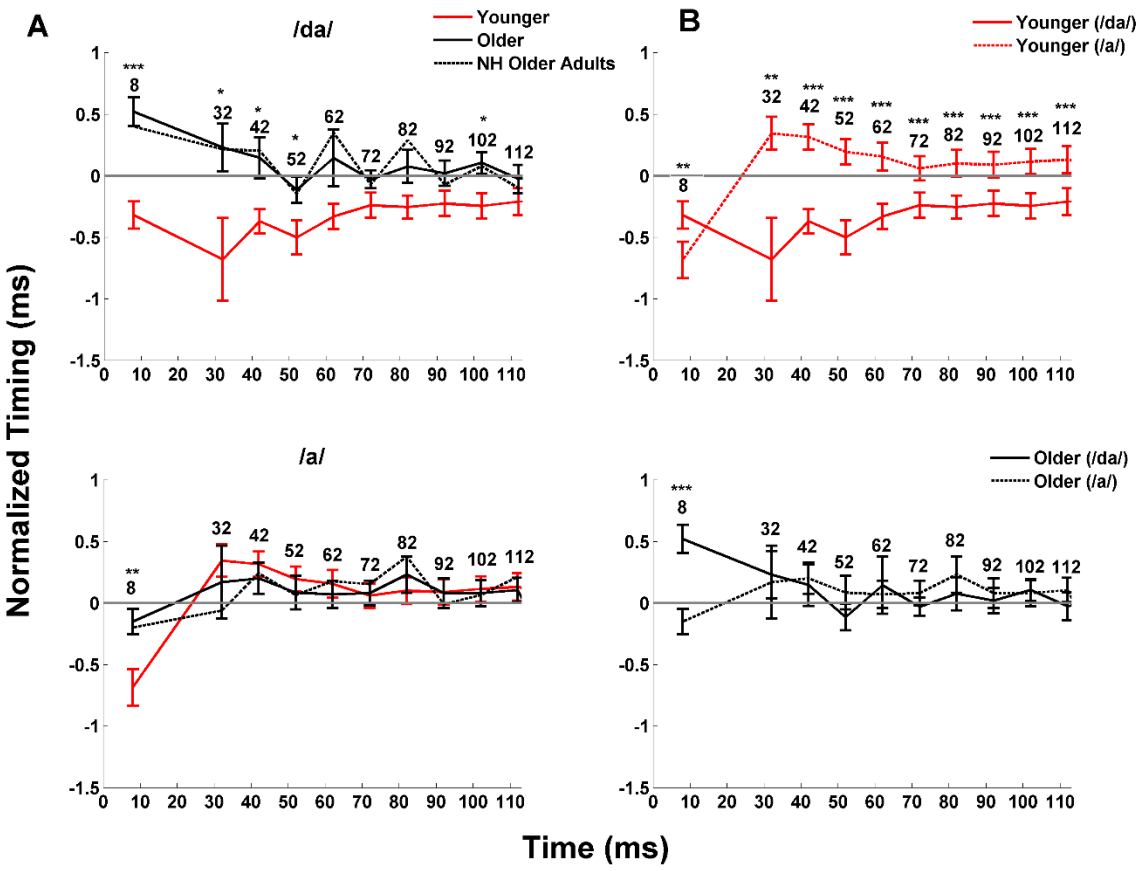

Fig. 14 Neural delays (mean ± 1SE) in the aging population for the /da/ and /a/. The x-axis represents the peak analyzed for each subject, while the y-axis represents the normalized peak latency for each subject. To facilitate visualization of the data, peak latencies on the y-axis were normalized. Normalization was obtained by subtracting the expected latency from the /da/ (8, 32, 42, 52, 62, etc.) from the actual response latency until 112 ms for the transition and the steady-state. For instance, if the latency detected at the onset for two different subjects were 7 ms and 9 ms, the correspondent normalized values would be -1 ms and 1 ms, respectively. Negative values indicate that the peaks were early with respect to the expected latency, while positive values indicate that the peaks were late with respect to the expected latency. A) In response to the /da/, older adults show a shift in neural response timing with respect to the younger adults for both the onset and transition (peaks 32-52) but not for the steady state, with the exception of peak 102. In response to the /a/, the only shift in neural response timing with respect to younger adults was observed at the onset. B) Latency difference within age groups. Older adults' latency response does not seem to be affected by the high frequency components of the speech syllable /da/ (all $p > 0.05$) with the exception for the onset. Conversely, younger adults' latency response show significant differences throughout the whole response (all $p < 0.01$) *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, N.S. = not significant.

Response and spectral magnitudes. Root mean square (RMS) amplitude was used to objectively quantify the overall magnitude of response and prestimulus (i.e., nonresponse) activity. RMS amplitudes were computed for the prestimulus period (-41 to 0 ms), the transition (18 - 68 ms), steady-state I (SS-I = 68 - 117 ms), steady-state II (SS-II = 117 -

169) and for the offset response (169 - 188 ms). The steady-state region was divided into two sections because of an observed drop in amplitude in the older adults' responses after ~115 ms. Average spectral amplitudes were calculated from each response using the fast Fourier transform (FFT) with zero padding and 1 Hz resolution over the two steady-state time regions (68-117 ms and 117-169 ms) in 40 Hz bins for the F_0 and second and third harmonics. FFT was calculated using the Welch method (50% overlapping). The average spectral amplitudes were used for statistical analysis, while the power spectral density (psd) was converted in dB/Hz and used to plot the frequency response.

Time-Frequency Analysis. Wavelets were used to analyze the grand-average of the responses of the envelope in the time–frequency domain. The complex Morlet's wavelets, $w(t, f_0)$, were used to decompose the signal between 80 and 1000 Hz (Tallon-Baudry et al. 1997). Morlet wavelets expressed as $w(t, f_0) = A\exp(-t^2/2\sigma_t^2)\exp(2i\pi f_0 t)$ have Gaussian shapes both in time and frequency domains with f_0 denoting central frequency. Time and frequency standard deviations σ_t and σ_f are related to each other as $\sigma_f = 1/(2\pi\sigma_t)$ and the factor $A = (\sigma_t\sqrt{\pi})^{-1/2}$ is used to normalize the total energy of the wavelets to 1. This wavelet family is characterized by $f_0/\sigma_f = 7$. The time-varying amplitude was calculated by taking the absolute values of the convolution of the complex wavelet with the signal: $Ampl(t, f_0) = |w(t, f_0) \times s(t)|$.

PLV Analysis. The PLV in the 80 – 1000 Hz range of the grand averages of younger versus older adults for the /a/ and /da/ was calculated using the same procedure and equations as

reported by Lachaux et al. (2002). In contrast with the phase-locking analysis performed by Anderson et al. (2012), this PLV was not applied to single trials but rather to the group average waveforms of younger and older participants. This analysis was used to evaluate the level of phase synchronization between responses of younger and older adults to /da/ and /a/. The range of values of the PLV goes from 0 (absence of phase coherence between the two signals analyzed) to 1 (perfect phase coherence between the two signals analyzed). The advantage of using this mathematical model is that only the phase of each continuous wavelet transform contributes to the final value of the phase coherence, as the amplitude has been removed. Moreover, wavelets ensure an optimal compromise time-frequency resolution for our frequency range.

Statistical analyses. All statistical analyses were conducted in IBM SPSS statistics software, version 21.0. A doubly multivariate repeated measures analysis of variance (ANOVA) was used to test group differences between transition and steady-state regions and between /a/ and /da/ in the time domain. This design was adopted because it permits analyses in which the dependent variables represent measurements of several variables for the different levels of the within-subjects factors. One-way analyses of variance were used for peak latencies (younger vs. older) and for RMS amplitudes, the F_0, and its harmonics. Levene's test was used to ensure homogeneity of variance for all measures. Paired t-tests were used to test RMS amplitude and latency differences within age groups in the time domain. Because the requirements for the Levene's test were not met for all our analyses, the non-parametric Mann-Whitney U and Friedman tests were applied to replace one-way ANOVA and one-way ANOVA with repeated measures, respectively.

RESULTS

Timing

The timing analysis partially refuted our hypothesis that older adults have impaired auditory processing for dynamic but not static regions of the speech stimuli used in this experiment. Overall, we found that the age-related differences in the transition region of the /da/ were primarily due to reduced encoding of the high-frequency energy of the /d/ consonant rather than to impaired temporal processing. This conclusion is based on the following analyses: we first tested differences between transition and steady-state regions in the two syllables between younger and older adults by using a doubly multivariate repeated measures model with 4 levels (/da/ vs. /a/, transition vs. steady state regions) and 3 peaks (32 ms, 42 ms and 52 ms peaks in the transition or 72 ms, 82 ms and 92 ms in the steady-state regions). Results indicated a significant group × region × syllable interaction (F[1,26]= 3.171, p = 0.041). This interaction was driven by both group × region (F[1,26] = 3.869, p = 0.021) and region × syllable interactions (F[1,26] = 4.38, p = 0.011).

To determine the factors that led to these interactions, we performed follow-up between-group and within-group testing. Follow-up between-group analyses for peak latencies were performed by using one-way ANOVAs (Figure 14 A). Younger adults had earlier latencies than older adults for the onset of both the /a/ and /da/ stimuli (F[1,28] = 8.737, p = 0.006 and F[1,28] = 27.636, p < 0.001, respectively). Results showed a group latency difference that is trending towards significance in the transition region in the /da/ (F[1,28] = 8.014, p = 0.064), but not in /a/ (F[1,28] = 0.162, p = 0.921), and no significant group latency differences in the steady-state region for either the /da/ (F[1,23] = 2.150, p = 0.0860 or the /a/ (F[1,28] = 0.381, p = 0.629).

Although these results might support the hypothesis that impaired temporal processing resulted in delayed latencies for the dynamic region of the syllable, the follow-up within-group testing suggested a primarily peripheral cause. Within-group testing showed earlier latencies for the /da/ than for the a/ in the younger group for both the transition (F[1,12] = 22.678, p < 0.001) and steady-state regions (F[1,12] = 18.761, p < 0.001), as expected given cochlear tonotopicity. However, the /da/ latencies were not earlier than the /a/ latencies in the older group for either the transition (F[1,12] = 0.970, p = 0.439) or the steady-state regions (F[1,12] = 0.618, p = 0.616), suggesting that reduced audibility from damage to the basal end of the cochlea resulted in latency delays for the /da/ (Figure 14 B). The expected time delay between syllables was only observed in older adults at the onset of the stimulus (t[14] = 6.686 p < 0.001, paired t-test).

The click analysis suggested that the temporal processing deficits might also be a contributing factor to delayed latencies. The click-evoked auditory brainstem response is dominated by contributions from the basal end of the cochlea (Don and Eggermont 1978). The latencies for Wave I should be affected by the audibility of the click, yet no significant differences were found in wave I (F[1,28] = 1.036, p = 0.317). On the contrary, click latencies for wave V were significantly different between groups (F[1,28] = 11.48, p = 0.002). Therefore, although audibility appears to be a primary factor for the age-related latency delays for the /da/, we cannot rule out the possibility of central contributions to these delays.

Effects of high frequency hearing loss on latency. Given the presence of high-frequency hearing loss in many of our older adults, we repeated the analysis after removing the subjects whose absolute audiometric hearing threshold of the right and left ears at 6000

and 8000 Hz was greater than 25dB HL. One-way ANOVAs between younger adults (n = 15) and the subset of older adults (n = 6) also showed significant differences at the onset and at peaks 42 ms, 62 ms and 82 ms in response to /da/ (all p values < 0.05), while no significant differences were found in response to the /a/ (all p values > 0.05) as seen in Figure 4 A (black dashed lines). However, there was also a similar pattern for the between-syllable comparison within older adults, in whom significant differences between the two speech syllables were observed only at the onset (t[14] = 8, p < 0.001) and at peak 112 ms (t[14] = -3.882, p = 0.012). Therefore, age-related cochlear changes were a factor in the delayed /da/ latencies even in this subset of older adults with relatively better hearing.

Response magnitude: time and frequency domains

Time domain: One-way ANOVAs were applied to the RMS values in the pre-stimulus, transition, SS-I, SS-II (only in response to the /da/) and offset (only in response to the /da/) regions between age groups. Because Levene's test of equality of error of variance failed for SS-II and offset in response to /a/, the non-parametric Mann-Whitney U test was used for these two cases. Results represented in Figure 15A showed that the response amplitudes of the younger adults were significantly higher than the older adults in the SS-I and SS-II regions for both the /a/ (F[1,28] = 7.656, p = 0.009, U [28] = 22, Z = -3.754, p < 0.001, respectively) and the /da/ (F[1,28] = 6.641, p = 0.015, F[1,28] = 8.979, p = 0.005, respectively), while no significance differences were found in the transition region (F[1,28] = 3.855, p = 0.059 for the /a/ and F[1,28] = 2.629, p = 0.116 for the /da/) and in the offset (U [28] = 82, Z = -1.265, p = 0.216 for the /a/ and F[1,28] = 2.696, p = 0.111 for the /da/).

A within-group comparison of RMS amplitude differences between the /da/ and /a/ syllables was consistent with our latency analysis. The older adults showed no significant amplitude differences between responses to /da/ and /a/ ($p > 0.05$, paired t-test) in any of the time regions except for SS-II ($t[14] = -3.557$ $p = 0.003$, paired t-test). Conversely, in younger adults significant differences were found in the transition and off-set regions ($t[14] = 2.353$ $p = 0.034$ and $t[14] = -3.848$ $p = 0.002$, respectively; paired t-test).

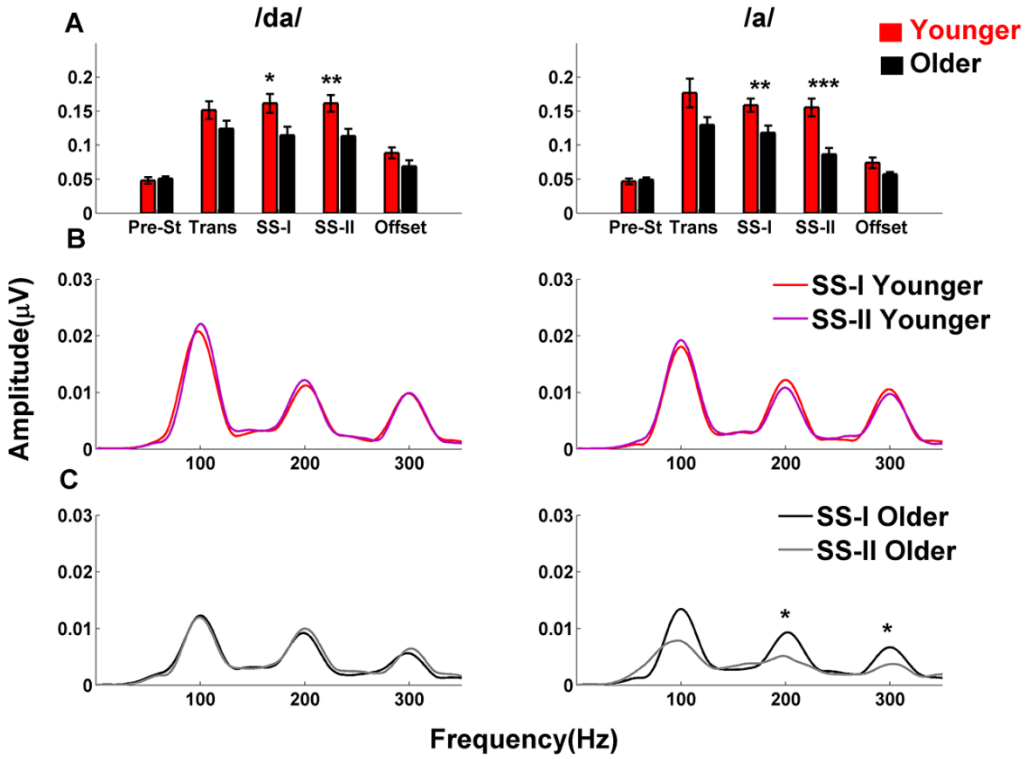

Fig. 15 A, Bar graphs showing the root mean square (RMS) values (mean ± 1 SE) for younger (red) and older (black) adults in response to /da/ and /a/. Older adults have lower RMS values in response to the steady-state regions in both the /da/ and /a/. B and C, Fast Fourier transforms calculated for the steady state (SS)-I and SS-II regions for the younger (B) and older (C) participants. Solid lines (red for younger and black for older adults) represent the spectral amplitudes for the SS-I, while dashed lines (green for younger and gray for older adults) represent spectral amplitudes for the SS-II. Note the difference between the SS-I and SS-II regions in older adults in response to the /a/ stimulus, only. *$p < 0.05$, **$p < 0.01$, ***$p < 0.01$.

Frequency domain: One-way ANOVAs were used to investigate differences in spectral amplitudes between younger and older adults in response to /da/ in the SS-I and SS-II

regions and in response to /a/ in the SS-I region. The Mann-Whitney U test was used for the response to /a/ in the SS-II region. Results indicated that younger adults had significantly higher amplitudes than older adults in the F_0 and in the third harmonic (H_3), but not in the second harmonic (H_2) in response to /da/ in the SS-I (F_0: $F[1,28] = 4.523$, $p = 0.042$; H_2: $F[1,28] = 1.641$, $p = 0.211$; H_3: $F[1,28] = 8.302$, $p = 0.007$) and SS-II (F_0: $p = 0.007$; H_2:$p = 0.412$; H_3:$F_{(1,28)} = 7.609$, $p = 0.010$) regions. Significantly higher amplitudes in response to /a/ were found in younger adults only in the H_3in the SS-I region (F_0: $F[1,28] = 2.248$, $p = 0.145$; H_2: $F[1,28] = 3.817$, $p = 0.061$; H_3 $F[1,28] = 8.281$, $p = 0.007$) and in F_0 and the second and third harmonics in SS-II region (F_0: $U[28] = 44$, $Z = -2.841$, $p = 0.003$; H_2: $U[28] = 39$, $Z = -3.049$, $p = 0.001$; H_3: $U[28] = 38$, $Z = -3.09$, $p = 0.001$).

Within-group comparisons between the SS-I and SS-II regions showed that older adults had higher amplitudes in the SS-I compared to the SS-II regions for both the F_0 ($t[14] = 3.031$ $p = 0.009$) and for the first two harmonics ($t[14] = 3.231$ $p = 0.006$ and $t[14] = 2.877$ $p = 0.012$, for H_2 and H_3 respectively) in response to the /a/, while no significant differences were found in response to the /da/ (all p values > 0.05). No significant within-group differences between the SS-I and SS-II regions were found in the F_0 and in the harmonics in response to /da/ and /a/ in younger adults (all p values > 0.05) as seen in Figures 15B and C.

The percentage difference between the SS-I and SS-II regions for the F_0, H_2, and H_3 in both /da/ and /a/ were also calculated in younger and older adults. Given the fact that the requirements for the Levene's test were not met, the non-parametric Friedman test was used to estimate the following: 1) if the percentage differences in /da/ and /a/ were

significantly different within age groups, and 2) if the percentage differences in /da/ and /a/ were significantly different between age groups. Results reported in Table 2 showed that significant differences within groups (/da/ vs /a/) were found in older adults (χ^2_5 = 15.724, p = 0.007), but not in younger adults (χ^2_5 = 3.229, p = 0.664), while significant differences between groups were found in response to /a/ (χ^2_5 = 22.352, p < 0.001), but not to /da/ (χ^2_5 = 1.248, p = 0.940). Figure 16 shows the waveforms for each subject in response to /da/ and /a/.

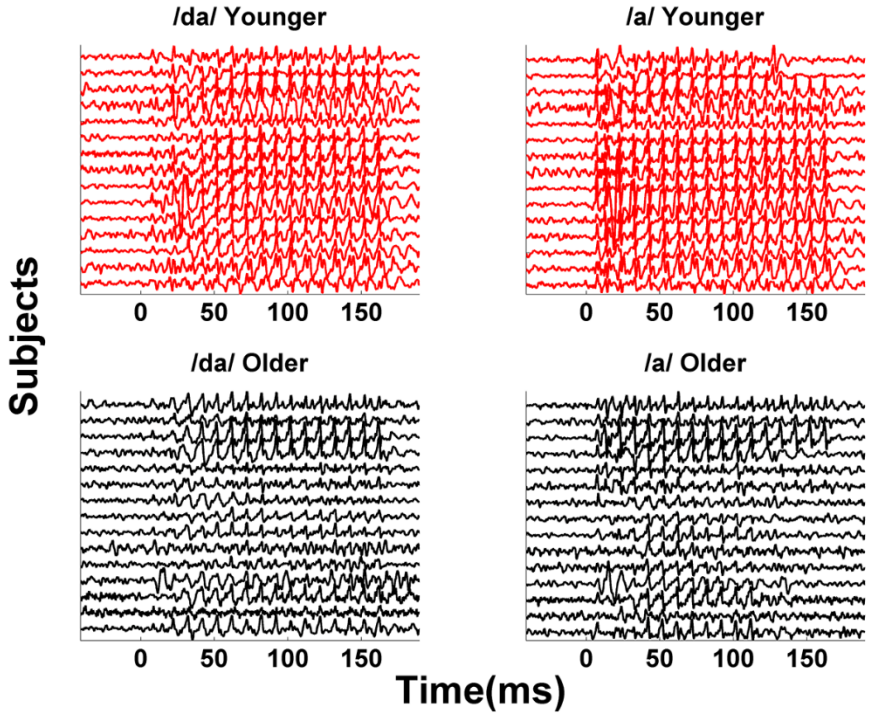

Fig. 16 Time series of for each subject in response to /da/ and /a/. Note how the steady state II tends to desynchronize in the majority of older adults. This phenomenon is observed only in response to /a/.

The percentage differences between the RMS values of the SS-I and SS-II regions were also calculated to estimate the effect of loss of synchronization for each subject. As in the case of the RMS values, because Levene test of equality of error of variance failed

for percentage difference for the SS-II region in response to /a/, the nonparametric Mann–Whitney *U* test was used. Results reported in Table 2 show that there was no significant difference between younger and older adults in response to /da/ [$F(1,28) = 0.051$, $p = 0.823$]. However, the loss of synchronization in SS-II in the majority of older adults in response to /a/ resulted in a dramatic decrease in amplitude, which was not observed in younger adults. This is reflected by significant differences between the two age groups [$U(28) = 45$, $Z = -2.8$, $p = 0.004$].

Table 2 Group means (with standard errors) for the younger and older adults for the percentage difference of the fundamental and of the first two harmonics and for the percentage difference of the root mean square values

	F **(/da/)**	***F*** **(/da/)**	***F*** **(/da/)**	***F*** **(/a/)**	***F*** **(/a/)**	***F*** **(/a/)**	**RMS (/da/)**	**RMS (/a/)**
Younger	−6.89 ± 7.14	−6.43 ± 6.68	−0.53 ± 7.53	6.47 ± 13.58	34.27 ± 18.75	38.66 ± 29.78	0.33 ± 2.57	5.95 ± 4.52
Older	9.16 ± 13.72	0.15 ± 9.28	−6.42 ± 7.68	74.24 ± 18.46	164.12 ± 61.18	122.01 ± 41.36	1.24 ± 3.14	40.18 ± 9.88

Time-Frequency analysis: The time-frequency analysis for the grand average of younger and older adults in response to /da/ and /a/ was carried out using Morlet wavelets. Results displayed in figure 17 suggest that higher harmonics (H_2 to H_4) in younger and older adults synchronize later to the /da/ than to the /a/. In response to the /da/, the younger and older adults synchronized at approximately 45 and 55 ms, respectively, and in response to the /a/, they synchronized at 10 and 25 ms, respectively. Synchronization to the steady-state region was preserved throughout the whole stimulation time in both age groups in response to the /da/ but faded around ~115 ms in the older adults only in response to the

/a/. The spectrogram also showed higher amplitude in younger adults at all frequencies for both /da/ and /a/.

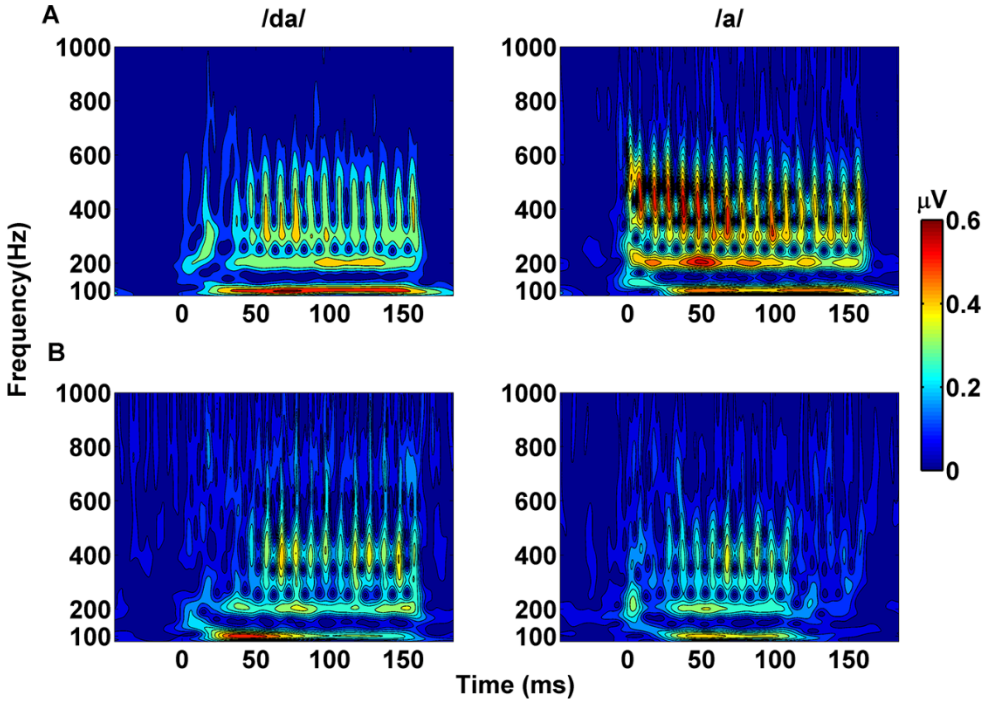

Fig. 17 Amplitude of the time–frequency analysis of the grand averages of the younger (A) and older populations (B). "0" indicates the stimulus onset. Note how the amplitude of the older population in response to /a/ drastically decreases at ~120 ms.

Phase-Locking Value: Results from the PLV analysis of the grand average are displayed in Figure 18, where differences in phase locking between the two age groups are shown. A significance threshold of ≥ 0.95 was chosen according to the values calculated by Lachaux et al. (2002). PLV values confirmed that the harmonics (H_2 to H_4) of older and younger adults synchronized earlier in response to the /a/ stimulus than in response to the /da/ stimulus. This is evident when looking at the first 40 ms, where robust synchronization between younger and older adults' responses in the first three harmonics (up to 400 Hz) was

observed in response to /a/, while a significant level of synchronization (black color) between the two age groups was not achieved until after 40 ms in response to / da/. This delay in synchronization to the /da/ is consistent with the peak latency delays in responses to /da/ resulting from reduced audibility. The synchronization to the /a/ was consistent up to ~115 ms, beyond which point the harmonics and the F_0 tended to desynchronize. Conversely, the synchronization in response to the /da/ was stable through 160 msec. These observations were consistent with both the frequency and the time-frequency analysis, where a clear reduction of amplitude was observed at around 115 ms in older adults in response to the /a/ and at around 160 ms in both younger and older adults in response to the /da/.

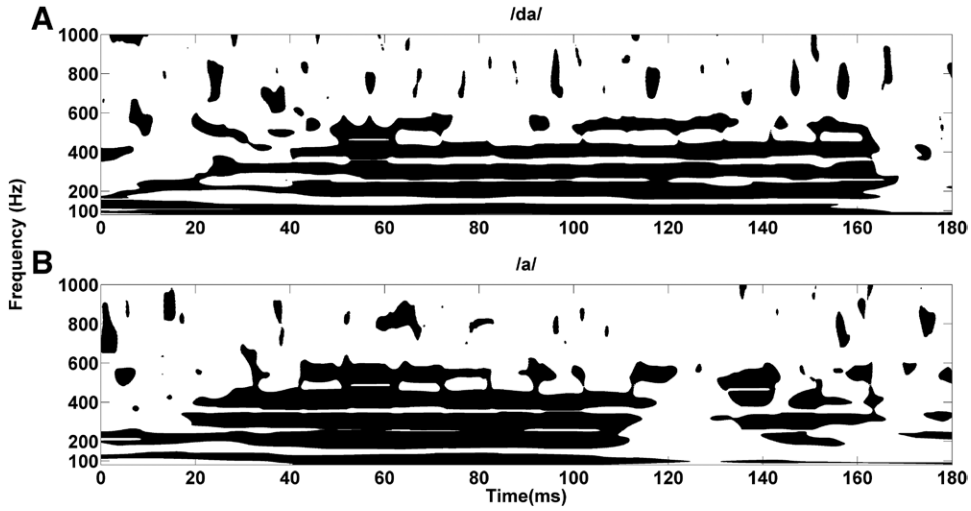

Fig. 18 Phase-locking value (PLV) for /da/ (younger vs. older) (A) and PLV for /a/ (younger vs. older) (B). The black color signifies that the responses of the younger and older adults are significant in phase (values ≥ 0.95), while the white color signifies that the responses of the two groups are out of phase (values < 0.95). The threshold value of 0.95 has been chosen based on the parameters used to calculate the PLV.

Discussion

The results did not entirely support the hypothesis that aging effects on subcortical responses are due to central rather than peripheral causes; rather, both peripheral and central factors appeared to contribute to the findings. The older adults' latency delays in

the CV transition were likely driven by reduced audibility (i.e., peripheral), while the loss of synchronization that was present in the late stage of the steady-state region in response to the speech-syllable /a/ in older adults may arise from impaired central processing.

Timing. Although we found the same latency delays for the transition region of the /da/ that were found in Anderson et al. (2012), our results suggest that these delays are at least partly due to a reduction of audibility for the high-frequency components of the /d/ consonant in older adults resulting from slight high-frequency hearing loss. There were significant group differences in hearing thresholds across the frequency range, with the greatest differences occurring at frequencies above 4 kHz. These hearing threshold differences likely account for the finding that responses of younger adults showed earlier peak latencies for the syllable /da/ than /a/ (as expected giving cochlear tonotopicity), while the same latency differences were not seen in the responses of older adults. Although the participants had clinically normal hearing, subclinical loss of outer hair cells may have contributed to the 1-msec difference in response latencies to the /da/ observed between younger and older adults. However, this mild high-frequency hearing loss did not result in delayed wave I click latencies in older adults compared with younger adults, while we did find age-related delays for wave V. This finding of delays for responses arising from the brainstem (wave V) but not from responses from the auditory nerve (wave I) provides evidence for a temporal processing deficit in addition to the delay caused by cochlear damage.

Response Magnitude: Time and Frequency Domain Static Versus Dynamic Regions. Aging affects response amplitudes both in the time and frequency domains, reflecting decreased encoding of the periodicity envelope derived from speech fine structure, consistent with what was previously reported using a similar 170-ms speech syllable /da/ (Anderson et al. 2012). Similar age effects, even though limited to the transition region, were found by Vander Werff and Burns (2011) and Clinard and Tremblay (2013), who used a 40-ms /da/. In contrast with the studies using a 40-ms /da/, the amplitude differences between younger and older adults reached significant values in the steady-state region, but not in the transition and offset regions, for both speech syllables. Because amplitude and latency should both be affected by temporal precision, these results appear to suggest that different neural mechanisms may underlie these age-related changes. It is possible that peripheral changes, such as damage to the cochlea, may be primarily responsible for the time delay observed in the responses of older adults. However, a loss of temporal precision at higher levels in the auditory system, such as the midbrain, may be the most plausible explanation for a loss of synchronization in the late stage of the steady-state response. It is most likely that both peripheral and central factors contributed to our findings. A reduction in auditory nerve fibers and loss of neural synchrony would presumably affect both latency and amplitude of the entire response. Furthermore, decreased inhibitory neurotransmission and other age-related changes may also have an impact on subcortical auditory processing in the midbrain (Caspary et al. 2006; Caspary et al. 2005; Walton et al. 1997). One of the functions of inhibitory neurotransmitters is to sharpen neural responses to rapidly varying acoustic stimuli (reviewed in Caspary et al. (2008)); therefore, a reduction in inhibitory neurotransmission may lead to a timing deficit specific to the changing formant transition region of the speech syllable. Because of the

frequency differences in our stimuli, however, we are not able to make this determination based on our results.

Loss of Synchronization in SS-II. A different mechanism may be responsible for the loss of amplitude noted in the SS-II region of the response of older adults to /a/. Synchronization appears to significantly decrease after ~95 ms (corresponding to ~115 ms from stimulus onset) of sustained phase locking, and this decrease is evident both in the time and frequency domains. Although surprising, this finding may arise from a loss of auditory nerve fibers that leads to an inability to sustain neural firing, such as may be found in abnormal acoustic reflex decay or tone decay test findings associated with VIIIth nerve lesions (Liden and Korsan-Bengtsen 1973; Olsen et al. 1975). Another possible explanation for the older adults' inability to sustain encoding a stimulus as efficiently as younger adults is prolonged neural refraction and loss of temporal synchronization among the neurons devoted to encoding that particular acoustic stimulus, as suggested by Walton et al. (1998). In that study, the integrity of temporal processing in the auditory midbrain was investigated by comparing detection of brief silent intervals in younger versus older CBA mice. The number of IC neurons that encoded short gap durations was reduced by approximately 50% in older mice. Furthermore, older mice had slower neural recovery times after previous stimulation with respect to younger mice, leading to poorer performance in detecting silent gaps. This ~50% reduction of IC neurons devoted to the encoding of short gap durations may be related in part to the loss of fibers with low spontaneous discharge rates from noise-induced cochlear neuropathy reported by Furman et al. (2013). Furman et al. speculated that this significant reduction in nerve fibers may cause hyperactivity that contributes to difficulty in processing auditory information in noisy environments by decreasing the signal-to-noise ratio. This effect was not found in response to

/da/, possibly because its steady-state region was shorter in duration. Our results seem to contrast with those of Bidelman et al. (2014b), who used a speech vowel continuum that did not include a transition region (similar to the /a/ used in our experiment) and did not report loss of synchronization. However, the length of the stimulus in that experiment was only 100 ms, and the effects observed in our study occurred after approximately 112 ms. Further studies that extend the steady-state region of the /da/ need to be conducted to explore the underlying mechanisms of this phenomenon.

Time-frequency analysis and PLV. Time-frequency analysis was applied to the grand averages of the younger and older adults to investigate representation of the fundamental frequency (100 Hz) and its harmonics. The results differed for the two stimuli. As expected, both younger and older adults showed a delay in synchronization in response to /da/ because of the presence of the stop consonant /d/. This delay was represented by smaller amplitudes at all frequencies in the onset region. However, in younger adults, the response became robust at all frequencies (higher amplitude) earlier (~30 ms) than in older adults (~46 ms). Consistent with what was observed in the frequency and time-domain analyses and with previous studies (Anderson et al. 2012; Clinard et al. 2010), differences in amplitude between the two age groups (Fig. 17) were present throughout the entire time region of stimulation and in all harmonics, while phase differences (Fig. 18A) were present in the transition region only. These phase differences observed in the transition region could be explained by the high-frequency components present in the consonant /d/ that older adults might have failed to encode due to cochlear damage. Failure to encode the high-frequency components of the speech syllable might have caused a delay in the response to the transition rather than poorer

phase locking to the stimulus. Conversely, in response to the /a/ stimulus, both younger and older adults show rapid synchronization, possibly because of the absence of high-frequency components present in the dynamic region of the /da/ (~20 and ~30 ms, respectively). Differences in amplitude were still evident throughout the entire stimulation time, suggesting that aging effects on amplitudes in midbrain were not limited to the dynamic encoding of speech. However, as observed in the time and frequency analysis, there was a drastic and significant loss of activity in the later region of the steady state of older adults, which was denoted by loss of amplitude (Fig. 17) and phase synchronization (Fig. 18B) in the last ~50 ms.

Limitations. We did not design this study to evaluate aging effects on sustained phase locking to a steady-state auditory signal. Therefore, a limitation of this study is our inability to elucidate the mechanism that underlies the sudden loss of synchronization in the SS-II region in response to the /a/. Furthermore, because of the higher frequency energy in the /da/ that is not present in the /a/, we were unable to rule out a peripheral contribution to the latency delays. To further investigate these phenomena, a follow-up experiment should compare responses to /a/ with responses to a CV speech syllable that has identical starting formants and steady-state regions of equivalent durations. In addition, it would be useful to include a group of older adults with significant hearing loss to further determine the extent to which loss of audibility contributes to latency delays.

Conclusion

Altogether, our findings of delayed latencies and reduced ability to sustain steady-state activity in older adults may be in line with results obtained with histological studies and single-neuron and single-fiber recordings. Animal models have shown that cochlear hair cell loss, decrease in synchronization, and prolonged neural refraction result in poor performance in tasks that involve rapid changes of the stimulus, such as in short gap detection tasks (Altschuler et al. 2015; Walton et al. 1998). These age-related changes may contribute to the older adult's difficulty with understanding someone who is talking rapidly or following a conversation in which there are multiple speakers. The results are in agreement with several studies that reported age-related changes in the peripheral and central auditory systems that affect the performance of older humans and animals in tasks in which temporal processing plays a critical role, such as when measuring auditory perception with distorted speech (Gordon-Salant and Fitzgibbons 1993) or variation of inter-onset intervals (Fitzgibbons and Gordon-Salant 2001) or when collecting physiological gap detection in a mouse model (Walton et al. 1997).

We have replicated the set of findings reported by Anderson et al. (2012) of diminished temporal precision in the FFRs in older adults who had normal hearing across the audiometric frequency range. Our study extends these findings and suggests that two different mechanisms could be responsible for the latency delay and loss of synchronization: cochlear damage and a loss of temporal precision in the midbrain. The decreased amplitude in the steady-state regions for both syllables and the sudden and unexpected loss of activity observed in the last 50 ms of the steady-state region in the older adults in response to the /a/ suggest that multiple mechanisms contribute to reduced

temporal precision and consequent impairments in speech perception. Aging affects auditory neural mechanisms in the form of loss of neurons and loss of synaptic connections (Willott 1996) and a decrease in inhibitory neurotransmission (Caspary et al. 2008). This impoverishment of neural networks may make the processing of some characteristics of sound harder to elaborate, such as tracking the CV transition or the sustained steady-state activity in a vowel. Imprecise representation of an auditory signal reduces the ability to selectively attend to that signal and extract meaning from it (Ding and Simon 2012). These findings advance understanding of the difficulties experienced by older adults when trying to understand speech in noisy conditions.

CHAPTER III

Second Study

Objectives

Humans have a remarkable ability to track and understand speech in unfavorable conditions, such as in background noise, but speech understanding in noise deteriorates with age. Results from several studies have shown that in younger adults, low frequency auditory cortical activity reliably synchronizes to the speech envelope, even when the background noise is considerably louder than the speech signal. However, cortical speech processing may be limited by age-related decreases in the precision of neural synchronization in the midbrain. To better understand the neural mechanisms contributing to impaired speech perception in older adults, we investigated how aging affects midbrain and cortical encoding of speech when presented in quiet and in the presence of a single competing talker. Our results suggest that central auditory temporal processing deficits in older adults manifest in both the midbrain and in the cortex. Specifically, midbrain frequency following responses to a speech syllable are more degraded in noise in older adults than in younger adults. This suggests a failure of the midbrain auditory mechanisms needed to compensate for the presence of a competing talker. Similarly, in cortical responses, older adults show larger reductions than younger adults in their ability to encode the speech envelope when a competing talker is added. Interestingly, older adults also showed an exaggerated cortical representation of speech in quiet conditions, suggesting a possible imbalance between inhibitory and excitatory processes that may impair their ability to efficiently encode speech.

INTRODUCTION

The ability to track and understand speech in the presence of interfering speakers is one of the most complex communication challenges experienced by humans. In a complex auditory scene, both humans and animals show an innate ability to detect and recognize individual auditory objects, an important component in the process of stream segregation. The ability to transform the noise-corrupted acoustic signal into a neural representation suitable for speech recognition may occur in the auditory cortex (AC) via adaptive neural encoding (Ding and Simon, 2012; 2013; Ding et al. 2014). Specifically, low frequency auditory cortical activity recorded with magnetoencephalography (MEG) reliably synchronizes to the slow temporal modulations of speech, even when the energy of the background noise is considerably higher than the speech signal, and even when the background noise is also speech. However, the accuracy of cortical speech processing may also be affected by the precision of neural synchronization in the auditory midbrain, as seen in studies that compare cortical responses to those using the frequency following response (FFR), believed to arise primarily from the midbrain (Chandrasekaran and Kraus 2010). For example, noise has a greater impact on the robustness of cortical speech processing in children (with learning impairments) who have delayed peak latencies in FFRs to a speech syllable (King et al. 2002). In young adults, earlier peak latencies in the FFR are associated with larger N1 amplitudes in cortical responses to speech in noise, and larger N1 amplitudes are associated with better ability to recognize sentences in noise (Parbery-Clark et al. 2011). Furthermore, Bidelman et al. (2014) demonstrated that temporal speech-processing deficits arising from the midbrain may be compensated by a stronger cortical response. Recent work from Chambers et al. (2016) showed that profound cochlear denervation can result in absence of auditory-evoked responses in the brainstem, but not in

the cortex, suggesting compensatory central gain increases that help restore the representation of the auditory object in auditory cortex. While these studies examined age- and hearing-loss related changes in midbrain and cortical encoding of vowels and tones presented in quiet, the comparison between midbrain and cortical encoding of speech syllables and sentences presented in competing single-talker speech has not yet been investigated in either younger or older adults.

Such auditory temporal processing deficits are of great relevance, since communication difficulties for older adults have a significant social impact, with strong correlations seen between hearing loss and depression (Carabellese et al. 1993; Herbst and Humphrey 1980; Kay et al. 1964; Laforge et al. 1992) and cognitive impairment (Gates et al. 1996; Lin et al. 2013; Uhlmann et al. 1989). Although audibility is an important factor in the older adult's ability to understand speech (Humes and Christopherson 1991; Humes and Roberts 1990), the use of hearing aids often does not improve speech understanding in noise, perhaps because increased audibility cannot restore temporal precision degraded by aging. Several electrophysiological studies in humans and animals support the hypothesis that degradation in central auditory temporal processes could play a role in explaining speech-in-noise problems experienced by older adults (Alain et al. 2014; Anderson et al. 2012; Clinard and Tremblay 2013; Lister et al. 2011; Parthasarathy and Bartlett 2011; Presacco et al. 2015; Ross et al. 2010; Soros et al. 2009).

To further investigate the neural mechanisms underlying age-related deficits in speech-in-noise understanding, this current study evaluated the effects of aging on temporal synchronization of speech in the presence of a competing talker in both cortex and midbrain. To de-emphasize the effects of audibility, only clinically normal hearing

listeners were included in both the younger and older age groups. Two experiments were conducted: the first one was used as an exploratory study at only a single SNR (~0 dB). The goal was to use these results to design and lay out the hypotheses for a second study that would aim at investigating the effects of aging at different signal-to-noise ratios (SNRs). We posit several hypotheses. First, in responses arising from midbrain, we hypothesize that younger adults encode speech with greater neural fidelity, reflected by higher amplitude responses and lower jitter, than older adults when the signal is presented in quiet and in noise. This hypothesis was driven by the results of the above mentioned studies showing more robust and less jittered responses in quiet in younger adults (Anderson et al. 2012; Clinard and Tremblay 2013; Mamo et al. 2015; Presacco et al. 2015) and an age-related effect of noise (Parthasarathy et al. 2010). In contrast, for cortical responses, we hypothesize that older adults will show an overrepresentation of the response both in quiet and noise. This hypothesis is driven by evidence showing age-related changes in both the amplitude (Alain et al. 2014; Soros et al. 2009) and the latency (Tremblay et al. 2003) of the main peaks of auditory cortical responses. Finally, we hypothesize that better speech-in-noise understanding (at the behavioral level) correlates with greater fidelity of neural encoding of speech, regardless of age.

MATERIALS AND METHODS

All procedures were reviewed and approved by the Institutional Review Board (IRB) of the University of Maryland. Participants gave informed consent and were paid for their time.

FIRST EXPERIMENT

Participants.

Participants for the midbrain-based EEG experiment comprised 15 younger adults (21 – 28 years, mean ± sd 23.13 ± 2.58, 4 males) and 15 older adults (60 – 76 years old, mean ± sd 64.66 ± 4.82, 6 males) recruited from the Maryland, Washington D.C. and Virginia areas. All participants had clinically normal hearing (Figure 19) defined as follows: (1) air conduction thresholds ≤ 25 dB HL from 125 to 4000 Hz bilaterally; and (2) no interaural asymmetry (> 15 dB HL difference at no more than two adjacent frequencies). Participants had normal IQ scores [≥ 85 on the Wechsler Abbreviated Scale of Intelligence (Zhu and Garcia 1999)] and were matched on IQ ($F_{(1,28)}$ = 2.008, p = 0.168) and sex (Fisher's exact, p = 0.472). In addition, the older adults were screened for dementia on the Montreal Cognitive Assessment (MOCA) (Nasreddine et al. 2005). All participants spoke English as their first language. The Quick Speech-in-Noise test (QuickSIN) (Killion et al. 2004), a commonly used clinical measure, was used to quantify the ability to understand speech in competing four-talker babble. Participants for the cortex-based MEG experiment comprised 8 younger (18 – 30 years, mean ± sd 23.87 ± 3.18, 3 males) and 8 older adults (60 – 68 years old, mean ± sd 63.37 ± 3.02, 3 males) who also participated in the EEG study. The smaller number of participants for the MEG study was because of several older adults being excluded due to magnetizable dental treatments. EEG and MEG data for each subject were collected in separate sessions.

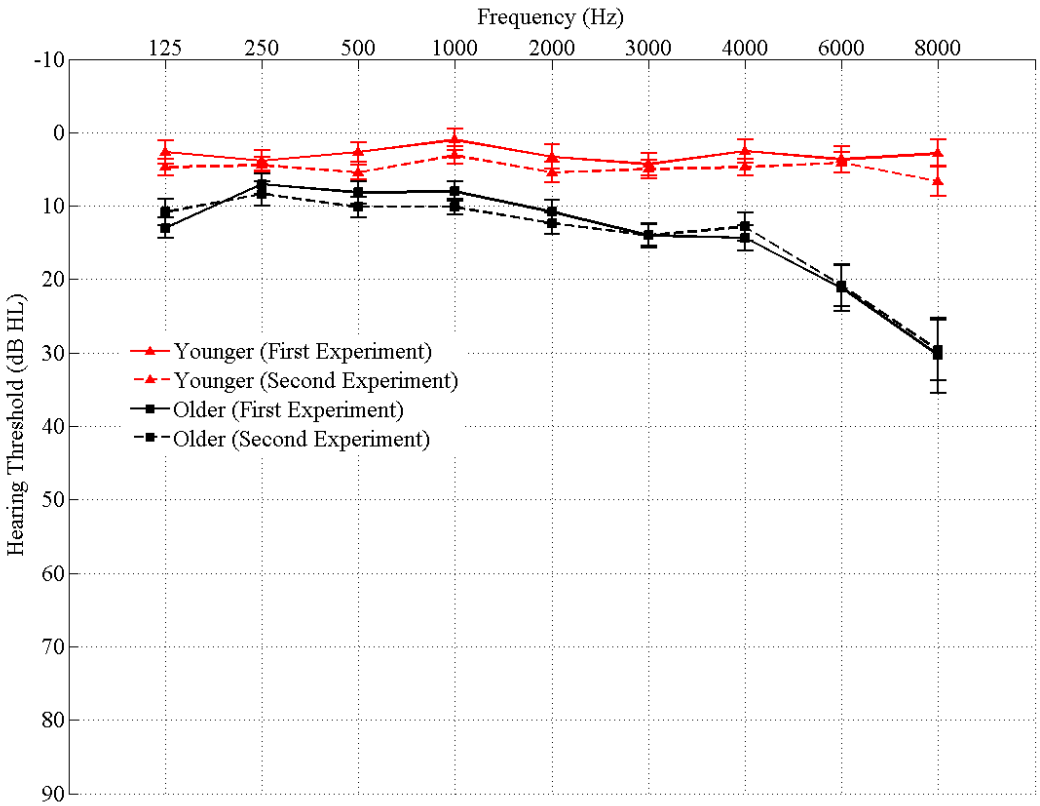

Fig. 19 Audiogram (mean ± 1SE) of the grand averages of both ears of younger (red) and older (black) adults for the first (solid line) and second (dashed line) experiment. All participants have clinically normal hearing.

EEG: Stimuli and recording.

A 170-ms /da/ (Anderson et al. 2012) was synthesized at a 20 kHz sampling rate with a Klatt-based synthesizer (Klatt 1980). The stimulus was presented at 80 peak dB SPL diotically with alternating polarities at a rate of 4 Hz through electromagnetically shielded insert earphones (ER-1; Etymotic Research) using Presentation software (Neurobehavioral Systems, Inc.). FFRs were recorded in quiet and in noise (0.45 dB SNR between the RMS values of the speech syllable /da/ and a single competing male talker used as the background noise) at a sampling frequency of 16384 Hz using the Biosemi Active Two acquisition system (BioSemi B.V.) with a standard vertical montage of 5 electrodes (Cz active, forehead ground, earlobe references). During the recording session (~1 hr),

participants sat in a recliner and watched a silent, captioned movie of their choice to facilitate a relaxed yet wakeful state. Three thousand artifact-free sweeps were recorded for each condition from each participant.

Data analysis.

Data recorded with Biosemi were analyzed in MATLAB (MathWorks, version R2011b) after being converted into MATLAB format with the function pop_biosig from EEGLab (Delorme and Makeig 2004). Sweeps with amplitude in the ±30 μV range were retained and averaged in real time and then processed off-line. The time window for each sweep was -47 to 189 ms referenced to the stimulus onset. Responses were digitally bandpass-filtered offline from 70 to 2000 Hz using a 4^{th} order Butterworth filter to minimize the effects of cortical low-frequency oscillations (Galbraith et al. 2000; Smith et al. 1975). A final average response was created by averaging the sweeps of both polarities to minimize the influence of cochlear microphonic and stimulus artifact on the response and simultaneously maximize the envelope response (Aiken and Picton 2008; Campbell et al. 2012; Gorga et al. 1985). Root-mean-square (RMS) values were calculated for the transition (18 ms to 68 ms) and steady-state (68 ms to 170 ms) regions. Correlation (Pearson's linear correlation) between the envelope response in quiet and noise was calculated for each subject to estimate the extent to which noise affects the FFR.

MEG recording.

Sixteen (eight per age group) of the 30 participants recruited for the EEG study were used for the MEG experiment. Participants were asked to attend to one of two stories

(foreground) presented diotically while ignoring the other story (background). The stimuli consist of segments from the book *A Child's History of England* by Charles Dickens, narrated by two different readers (of opposite gender). Each speech mixture was constructed as described by Ding and Simon (2012) by digitally mixing two speech segments into a single channel, with duration of 1 minute. Three trials of each stimulus were presented alternately, so that the first trial of stimulus 1 was followed by the first trial of stimulus 2 and so on. The sound recording was presented diotically at an intensity level of ~62 dB SPL using Presentation (Neurobehavioral Systems). The sounds (approximately 65 dB SPL when presented in noise) were delivered to the participants' ears with 50 Ω sound tubing (E-A-RTONE 3A; Etymotic Research), attached to E-A-RLINK foam plugs inserted into the ear canal. The entire acoustic delivery system was equalized to give an approximately flat transfer function from 40 to 3000 Hz, thereby encompassing the range of the presently delivered stimuli. At the end of each trial, participants were asked to answer a question about the story content to confirm they executed the task correctly and attended to the target speaker. Neuromagnetic signals were recorded using a 157-signal whole head MEG system (Kanazawa Institute of Technology, Kanazawa, Japan) in a magnetically shielded room with the same set-up used by Ding and Simon (2012).

Data analysis.

Three reference channels were used to measure and cancel the environmental magnetic field by using TS-PCA (De Cheveigné and Simon 2007). MEG data were analyzed off-line using MATLAB. The 157 raw MEG data channel responses were first filtered between 2 - 8 Hz with an order 700 windowed (Hamming) linear-phase FIR filter, then decomposed

into N signal components (where $N \leq 157$) using the denoising source separation (DSS) algorithm (de Cheveigne and Simon 2008; Särelä and Valpola 2005). The filtering range of 2 – 8 Hz was chosen based on previous results showing the absence of intertrial coherence above 8 Hz (Ding and Simon 2013) and the importance of the integrity of the modulation spectrum above 1 Hz to understand spoken language (Greenberg and Takayuki 2004). The signal components were extracted for each trial, band-pass filtered between 1 – 8 Hz (Ding and Simon 2012) with a 2nd order Butterworth filter, and averaged over trials. Reconstruction of the envelope was performed using a linear reconstruction matrix estimated via the Boosting algorithm (David et al. 2007; Ding et al. 2013; Ding and Simon 2013). Success of the reconstruction is measured by the linear correlation between the reconstructed and actual speech envelope. The reconstructed envelope was obtained from the speech of the single speaker to which the participant was instructed to attend only, not from the actual acoustic stimulus. The envelope was computed as the 1 - 8 Hz band-pass filtered magnitude of the analytic signal. Data were analyzed using 3 different time windows for this reconstruction model: 500, 350, and 150 ms. The choice to narrow the integration window down to 150 ms is based on previous results showing that the ability to track the speech envelope substantially worsens as the window decreases down to 100 ms (Ding and Simon 2013). These values refer to the time shift imposed on our data with respect to the onset of the speech and to the corresponding integration window of our reconstruction matrix. Specifically, if processing time for younger and older adults is the same, then their performance should follow the same pattern as the integration window changes. Conversely, if older adults require more time to process the information because of the possible presence of temporal processing deficits, a narrower integration window

should then negatively affect their performance more than for younger adults. The noise floor was calculated by using the neural response recorded from each condition to reconstruct the speech envelope of a different stimulus than was used during this response.

Statistical analyses.

All statistical analyses were conducted in SPSS version 21.0 (SPSS). Split plot ANOVA were used to test for age group × condition interactions for the RMS values of the FFR response in the time domain and for correlation values calculated for the cortical data. The Greenhouse-Geisser test was used when the Mauchly's sphericity test was violated. A paired t-test was used for within-subject group analysis for the correlation values and amplitudes for the cortical data, while one-way analyses of variance (ANOVAs) were used to analyze the RMS amplitude values of the FFR and the correlation values and the amplitude for the cortical data. The non-parametric Mann-Whitney U test was used in place of the one-way ANOVA when Levene's test of Equality of Variances was violated. Two-tailed Spearman's rank correlation (ρ) were used to evaluate the relationships among speech-in-noise scores, midbrain and cortical parameters. Fisher's z transformation was applied to all the correlation values calculated for the midbrain and cortical analysis. The false discovery rate (FDR) procedure (Benjamini and Hochberg 1995) was applied to control for multiple comparisons where appropriate.

SECOND EXPERIMENT

Participants.

Participants comprised 17 younger adults (18 – 27 years, mean ± sd 22.23 ± 2.27, 3 male) and 15 older adults (61 – 73 years old, mean ± sd 65.06 ± 3.30, 5 males) recruited from the Maryland, Washington D.C. and Virginia areas. Eight older and three younger adults that participated in the first study came back for the second experiment. All participants had clinically normal hearing (Figure 1) as defined in the first experiment. Participants had normal IQ scores [≥ 85 on the Wechsler Abbreviated Scale of Intelligence (Zhu and Garcia 1999)] and were matched on IQ ($F_{(1,30)} = 0.660$, $p = 0.423$) and sex (Fisher's exact, $p > 0.05$). In addition, the older adults were screened for dementia on the Montreal Cognitive Assessment (MOCA) (Nasreddine et al. 2005). All participants spoke English as their first language. The Quick Speech-in-Noise test (QuickSIN) (Killion et al. 2004) was used to quantify the ability to understand speech in noise composed of four-talker babble. All participated in both the EEG and MEG study. EEG and MEG data for each subject were collected in two different sessions.

EEG: Stimuli and recording.

The same speech syllable /da/, and procedures, used for the first experiment were used for the second, in quiet and also with 4 different SNRs: +3 dB, 0 dB, -3 dB and -6 dB.

Data analysis.

The same signal processing techniques used to analyze EEG data in the first experiment were applied to the data collected in the second experiment.

MEG: Stimuli and recording.

Participants were asked to attend to one of two stories (foreground) presented diotically while ignoring the other one. The stimuli for the foreground consist of segments from the book *The Legend of Sleepy Hollow* by Washington Irving, while the stimuli for the background narrated by an English speaker consist of a segment from the book *A Christmas Carol* by Charles Dickens. The foreground was always spoken by a male talker, while the background story was always spoken by a female talker. Additional stimuli using a background narration in an unfamiliar language were also presented, but the responses to those stimuli are not analyzed here. Each speech mixture was constructed as described by Ding and Simon (2012) by digitally mixing two speech segments into a single channel, with duration of 1 minute. Five different conditions were recorded: quiet, +3 dB, 0 dB, -3 dB and -6 dB SNR. Four different segments from the same foreground story were used in order to minimize the possibility that the clarity of the stories could affect the performance of the subjects. In order to maximize the level of attention of the subject on the foreground segment, participants were asked beforehand to count the number of times a specific word or name was mentioned in the story. The sounds (approximately 70 dB SPL when presented with a solo speaker) were delivered to the participants' ears with 50 Ω sound tubing (E-A-RTONE 3A; Etymotic Research), attached to E-A-RLINK foam plugs inserted into the ear canal. The entire acoustic delivery and MEG systems used were the same as described in the first experiment. The noise floor was calculated in the same way as described in the first experiment.

Statistical analyses.

All statistical analyses were conducted in SPSS version 21.0 (SPSS). The same statistical analyses used for the first experiment were used for the second experiment.

RESULTS

FIRST EXPERIMENT

Speech Intelligibility.

Younger adults (mean ± std = -0.30 ± 1.18 dB SNR loss) scored significantly better ($F_{[1,28]}$ = 16.128, $p < 0.001$) than older adults (mean ± std = 1.16 ± 0.78 dB SNR loss), suggesting that older adults' performance in noise may decline compared to younger adults even when audiometric thresholds are clinically normal.

Midbrain (EEG).

Amplitude analysis.

Figure 20 shows the grand average of FFRs to the stimulus envelope of younger and older adults in quiet and in noise. The decrease in younger adults' response amplitudes between quiet and noise conditions is significantly larger than in older adults in the transition (RMS × age group interaction ($F_{[1,28]}$ = 6.339, p = 0.018)) but not in the steady-state region (RMS × age group interaction $F_{[1,28]}$ = 0.919, p = 0.346). A one-way ANOVA showed no significant differences between younger and older adults in either the transition ($F_{[1,28]}$ = 3.88, p = 0.059 and $F_{[1,28]}$ = 0.689, p = 0.414 in quiet and noise respectively) or in the steady-state ($F_{[1,28]}$ = 3.226, p = 0.083 and $F_{[1,28]}$ = 3.077, p = 0.09 in quiet and noise respectively) regions. Despite the absence of significant group differences in the transition regions in both quiet and noise conditions, the grand average shown in Figure 2 suggests

that the interaction is driven by larger age group differences in the quiet condition than in the noise condition. The follow-up results of a paired t-test suggest that noise significantly decreases response amplitude in both younger and older adults in both the transition (18 ms – 62 ms; $t_{[14]} = 12.43$, $p < 0.001$ for younger adults and $t_{[14]} = 5.125$, $p < 0.001$ for older adults) and steady state regions (62 ms – 170 ms; $t_{[14]} = 5.91$, $p < 0.001$ for younger adults and $t_{[14]} = 3.896$, $p = 0.002$ for older adults).

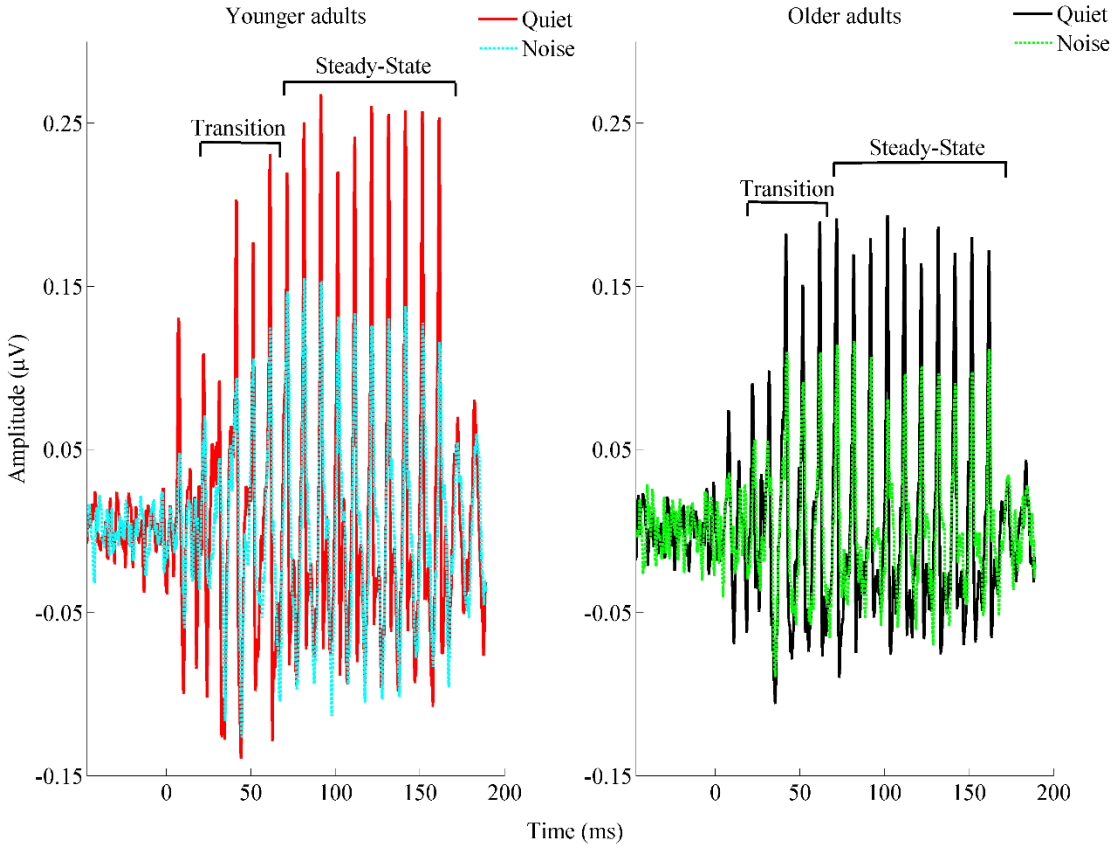

Fig. 20 Grand average (n = 15) of the response to the stimulus envelope for the two conditions of younger (left; quiet = red, noise = light blue) and older (Right; quiet = black, noise = green) adults. In the transition and steady-state regions, noise resulted in a significant decrease ($p < 0.001$) in the RMS amplitudes for both younger and older adults. No significant differences between groups were found in either quiet or noise. Statistical analysis showed a RMS × group interaction effect in the transition ($p = 0.018$), but not in the steady-state region ($p > 0.05$). The transition region is harder to encode for older than for younger adults.

Correlation analysis.

In order to analyze the robustness of the response in noise, we linearly correlated the average response obtained in quiet with the one obtained in noise, for both the transition and steady-state regions for each subject. Younger adults showed higher correlation values (less degradation in noise) than older adults, in both the transition and steady-state regions, but the differences between the two age groups did not reach significance in either region ($F_{[1,28]} = 2.000$, $p = 0.168$ for the transition and $F_{[1,28]} = 2.701$, $p = 0.111$ for the steady-state region; r-values were Fisher z transformed before statistical analysis).

Cortex (MEG): Reconstruction of the speech envelope.

The ability to reconstruct the low-frequency speech envelope from cortical activity is a measure of the fidelity of the neural representation of that speech envelope (Ding and Simon 2012). This reconstruction accuracy was compared between noise and quiet conditions. Figure 21 shows an example of reconstruction of the speech envelope of the foreground in noise from a younger and an older adult. Results from a split plot ANOVA applied to the three integration windows utilized for the analysis revealed a time window × age group interaction in quiet ($F_{[2,60]} = 5.037$, $p = 0.027$), but not in noise ($F_{[2,60]} = 1.090$, $p = 0.325$). Repeated measures ANOVA show that shortening the integration window from 500 ms down to 150 ms significantly impairs the reconstruction accuracy of only older adults in both quiet ($F_{[2,30]} = 20.623$, $p < 0.001$) and noise ($F_{[2,30]} = 7.090$, $p = 0.01$). Conversely, younger adults do not show any significant difference in either condition ($F_{[2,30]} = 4.161$, $p = 0.048$, and $F_{[2,30]} = 2.104$, $p = 0.158$, in quiet and noise respectively).

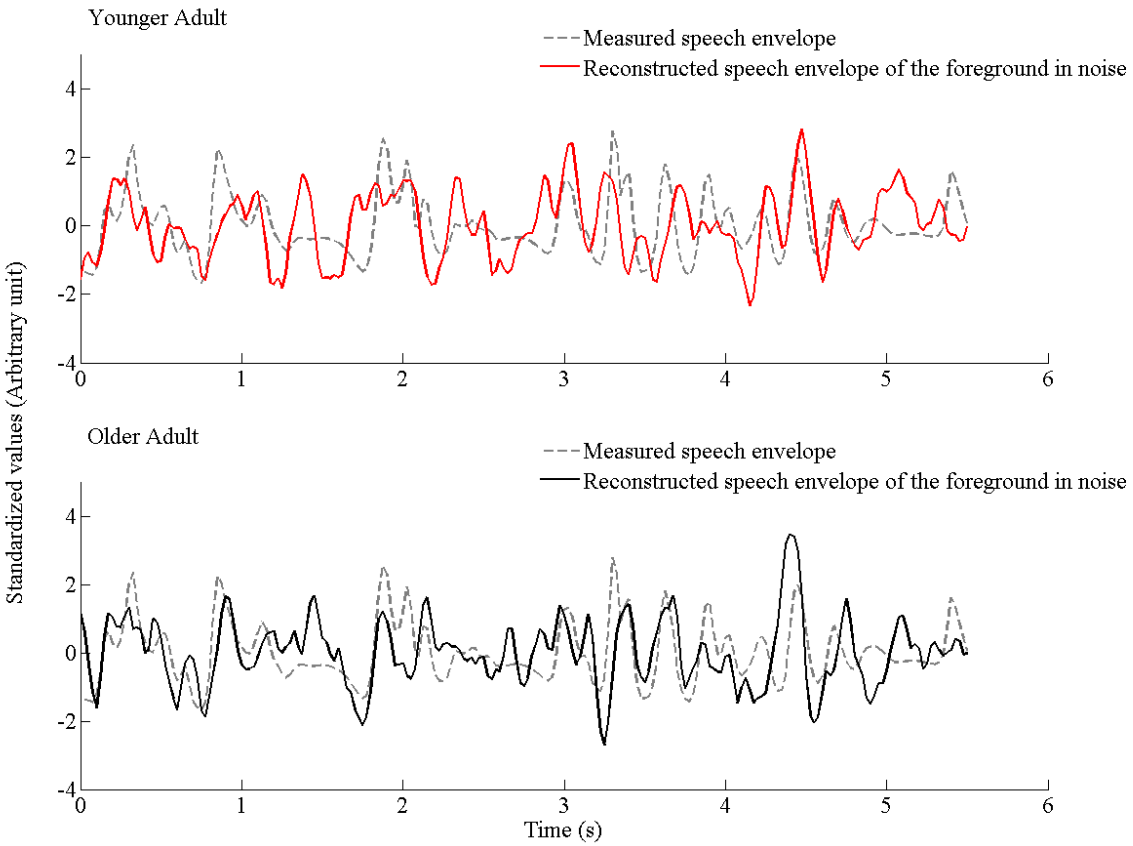

Fig. 21 Example of the reconstrcution of the speech envelope of the foreground for younger (top) and older (bottom) adults in noise.

A follow up paired t-test of the foreground reconstructed in quiet and in noise at 500 ms vs 350 ms and 150 ms showed that the reconstruction accuracy of younger adults is not significantly affected by the integration windows (t[15] = -1.164, p = 0.262 and t[15] = 1.1742, p = 0.102 in quiet for 500 ms vs 350 ms and 500 ms vs 150 ms and t[15] = -0.037, p = 0.971 and t[15] = 1.367, p = 0.192 in noise for 500 ms vs 350 ms and 500 ms vs 150 ms), while older adults' ability to track the speech envelope of the foreground is significantly reduced at 150 ms in both quiet (t[15] = -1.114, p = 0.283 and t[15] = 4.429, p < 0.001 500 ms vs 350 ms and 500 ms vs 150 ms, respectively) and noise (t[15] = 1.446, p = 0.169 and t[15] = 3.042, p = 0.008 for 500 ms vs 350 ms and 500 ms vs 150 ms, respectively). Interestingly, older adults also show a tendency to have greater fidelity of

the reconstruction at 500 ms in both quiet (mean ± 1SE, 0.23 ± 0.033 for older and 0.208 ± 0.021 for younger) and noise (mean ± 1SE, 0.204 ± 0.027 for older and 0.192 ± 0.025 for younger), even though this difference did not reach significance value in either conditions ($p > 0.05$). All the reconstruction values at 500 ms were significantly higher than the noise floor (all p values < 0.001). Figure 22 shows box plots of the mean reconstruction accuracy values in quiet and noise for younger and older adults.

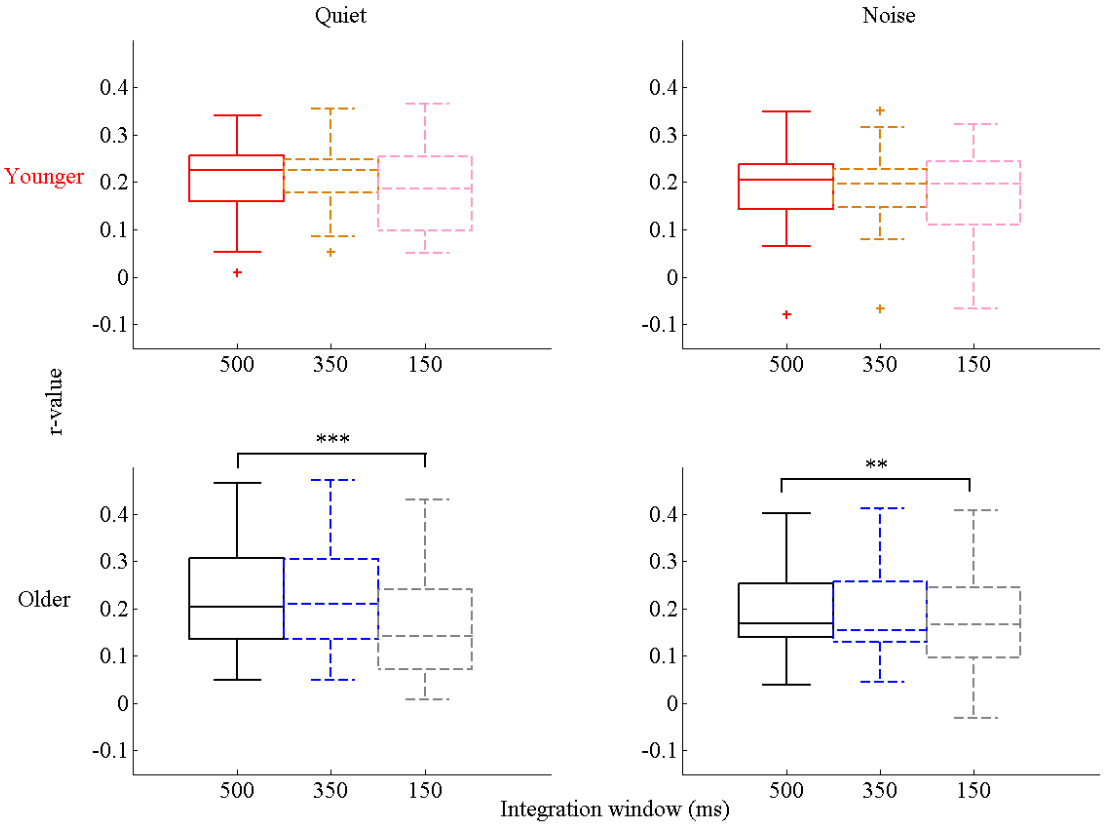

Fig. 22 Box plots (median, 25th and 75th percentile; whiskers extend to the most extreme data points not considered outliers, and outliers are plotted individually with the "+" symbol) representing the reconstruction accuracy of the foreground for younger (top) and older (bottom) adults in quiet and in noise for each of the 3 integration windows used in the analysis. There is significant drop in stimulus reconstruction accuracy when shortening the integration time available, but only for older listeners; for younger listeners even the shortest integration time is sufficient for reconstruction. ** $p < 0.01$, ***$p < 0.001$

Relationships among behavioral, midbrain and cortical data.

Two-tailed Spearman's rank correlation coefficient (ρ) was used to study the correlations among the following measurements: speech-in-noise score, cortical decoding accuracy in quiet and in noise, with an integration window of 500 ms, and the quiet-to-noise correlation value in the steady-state region of midbrain responses. No significant correlations were found in either younger or older adults in any of the relationships tested.

Brief remarks about the first experiment

Results from the first experiment suggest a trend towards greater fidelity of the midbrain response in younger adults, consistent with our hypotheses, but results were not significant in this first experiment. Conversely, an opposite tendency was found in the cortical data, which is that older adults had a higher speech reconstruction accuracy in both quiet and noise. Furthermore, their ability to track the foreground was significantly diminished, as the integration window was narrowed down to 150 ms. The same pattern was not observed in younger adults, suggesting that as we age the computational time required to perform a task might increase, as reported in several behavioral studies (Fitzgibbons and Gordon-Salant 2001; Gordon-Salant et al. 2006). Interestingly, no correlation between midbrain and cortex was observed, reinforcing the existence of compensatory central gain increases that could help restore the representation of the auditory object at the cortical level, even when the temporal processes in the midbrain are severely deteriorated (Chambers et al. 2016).

This experiment was limited by the use of a single SNR and a smaller number of qualified MEG subjects, which may have limited the possibility of separating midbrain and

cortical response patterns of younger and older adults and of finding any association between midbrain and cortical response patterns. Based on these limitations and observations, we designed a second experiment, using multiple SNRs and with a larger number of qualified MEG subjects.

SECOND EXPERIMENT

Speech Intelligibility.

Younger adults (mean ± std = -0.57 ± 1.13 dB SNR loss) scored significantly better ($F_{[1,30]}$ = 10.613, p = 0.003) than older adults (mean ± std = 0.8 ± 1.25 dB SNR loss), suggesting that older adults' performance in noise may decline compared to younger adults even when audiometric thresholds are clinically normal.

Midbrain (EEG).

Amplitude analysis.

Figure 23 displays the RMS values for each condition tested in younger and older adults in the transition and steady-state regions.

Transition region. A one-way ANOVA showed that younger adults have significantly higher RMS values in quiet ($F_{[1,30]}$ = 4.255, p = 0.048). When all the noise conditions were grouped together, one-way ANOVA showed significant differences between younger and older adults ($F_{[1,126]}$ = 5.150, p = 0.025). The follow-up results of paired t-tests suggest that noise significantly decreases response amplitude in both younger and older adults in all the noise conditions tested (all p values < 0.01). Repeated measures ANOVA showed a condition × age interaction between quiet and noise at -3 dB ($F_{[1,30]}$ = 6.264, p = 0.018) and

-6 dB ($F_{[1,30]} = 6.696$, $p = 0.015$), but not at the other conditions tested ($F_{[1,30]} = 1.125$, $p = 0.297$ and $F_{[1,30]} = 0.333$, $p = 0.568$, and for +3 dB and 0 dB respectively). Repeated measures ANOVA showed significant differences across noise conditions in younger ($F_{[3,48]} = 13.384$, $p < 0.001$), but not in older ($F_{[3,48]} = 0.885$, $p = 0.457$) adults.

Steady-state region. A one-way ANOVA showed that younger adults have significantly higher RMS values than older adults in quiet ($F_{[1,30]} = 6.877$, $p = 0.014$). The follow-up results of paired t-tests suggest that noise significantly decreases response amplitude in both younger and older adults in all the noise conditions tested (all p values < 0.05). Repeated measures ANOVA showed no condition × age interaction between quiet and meaningful noise at any of the conditions tested ($F_{[1,30]} = 0.072$, $p = 0.791$, $F_{[1,30]} = 0.000$, $p = 0.986$, $F_{[1,30]} = 2.574$, $p = 0.119$ and $F_{[1,30]} = 3.197$, $p = 0.084$ for +3 , 0, -3 and -6 dB respectively). Repeated measures ANOVA showed significant differences across noise conditions in younger ($F_{[3,48]} = 19.847$, $p < 0.001$), but not in older ($F_{[3,48]} = 0.874$, $p = 0.462$) adults. When all the noise conditions were grouped together, a follow-up one-way ANOVA showed significant differences between younger and older adults ($F_{[1,126]} = 27.364$, $p < 0.001$).

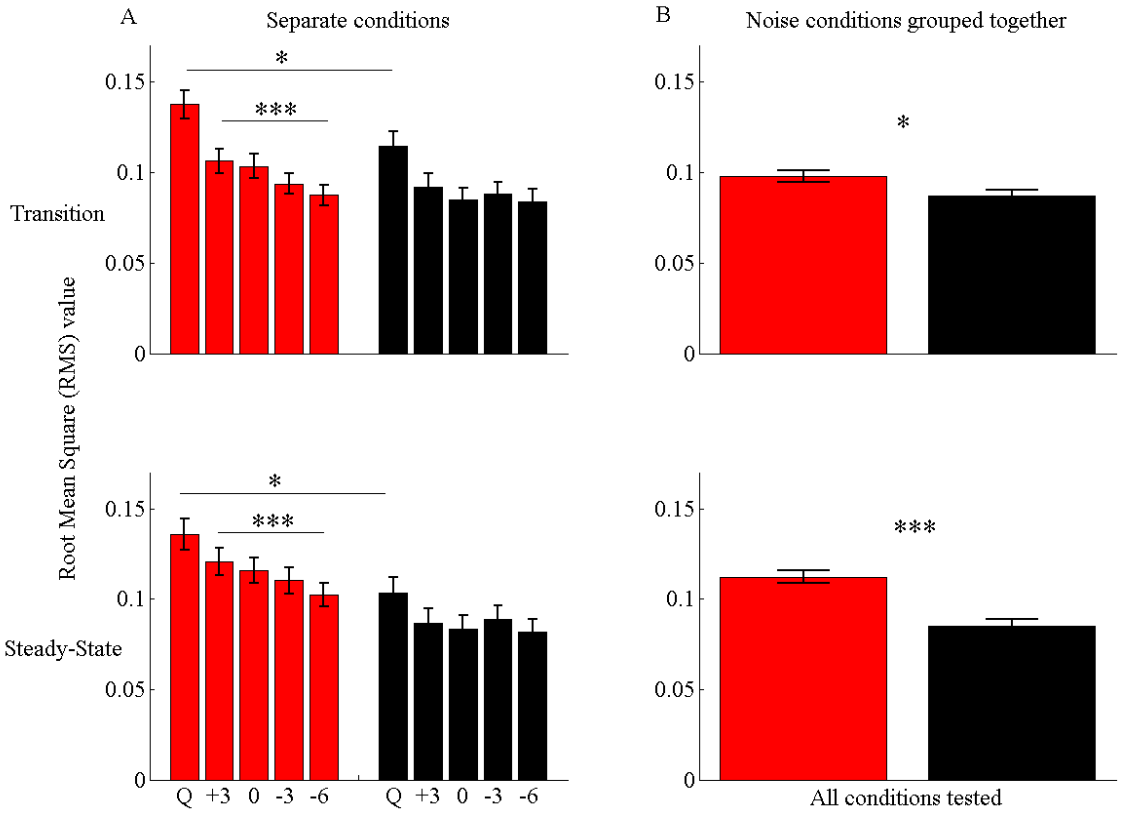

Fig. 23 RMS value ± 1SE of the envelope for the conditions (Q = Quiet, +3 = +3 dB, 0 = 0 dB, -3 = -3 dB and -6 = -6 dB) tested in younger (red) and older (black) adults. Column A represents the average RMS for each single condition, while column B shows the average RMS for all the noise conditions tested. Younger adults had significantly higher RMS values in quiet in both the transition and the steady-state regions. A RMS × group interaction effect was noted in the transition at -3 and -6 dB, but not in the steady-state region. Repeated measures ANOVA applied to the four noise conditions show significant differences in younger adults in both the transition and steady-state regions, but not in older adults. Noise minimally affects older adults, likely because their response in quiet is already degraded. *$p < 0.05$, ***$p < 0.001$

Correlation analysis. In order to analyze the robustness of the response in noise, we linearly correlated (Pearson correlation) the average response obtained in quiet with the ones recorded in noise, for both the transition and steady-state regions for each subject. Repeated measures ANOVA showed no significant noise condition × age interaction in either the transition ($F_{[3,90]} = 1.129$, $p = 0.342$) or the steady state ($F_{[3,90]} = 1.015$, $p = 0.390$) region. When all the noise conditions were grouped together, a follow-up Mann-Whitney U test showed significantly higher r-values in younger adults in the steady-state (*U* [128]

= 1272, Z = -3.667, p < 0.001), but not in the transition region (U [128] = 1675, Z = -1.743, p = 0.081).

Cortex: Reconstruction of the speech envelope. The ability to reconstruct the low-frequency speech envelope from cortical activity is a measure of the fidelity of the neural representation of that speech envelope (Ding and Simon 2012). Figure 24 displays the grand average ± standard error of the reconstruction accuracy for younger and older adults for all the conditions tested. A one-way ANOVA showed significantly higher correlation values in quiet in older adults ($F_{[1,30]}$ = 14.923, p = 0.001). A one-way ANCOVA (with correlation in quiet used as covariate) showed no significant differences between younger and older adults at any of the noise conditions tested ($F_{[1,29]}$ = 2.633, p = 0.115, $F_{[1,29]}$ = 1.906, p = 0.178, $F_{[1,29]}$ = 3.308 p = 0.079, $F_{[1,29]}$ = 0.005, p = 0.946 for +3 dB, 0 dB, -3 dB and -6 dB respectively). All the reconstruction values were significantly higher than noise floor (all p values < 0.01). Since the difference between older and younger adults is minimized at -6 dB, this condition was used to analyze the effect of the integration window on the fidelity of the reconstruction of the speech envelope. Results from a split plot ANOVA applied to the three integration windows utilized for the analysis revealed a reconstruction window × age group interaction in quiet ($F_{[2,60]}$ = 9.332, p = 0.004), but not in noise ($F_{[2,60]}$ = 0.105, p = 0.802). Repeated measures ANOVA applied to 500, 350 and 150 ms integration windows show significant differences in older adults in both quiet ($F_{[2,32]}$ = 14.954, p = 0.00) and noise ($F_{[2,32]}$ = 5.048, p = 0.037), but not in younger adults ($F_{[2,32]}$ = 4.213, p = 0.048 and $F_{[2,32]}$ = 1.195, p = 0.302, in quiet and noise respectively). A follow-up paired t-test of the foreground reconstructed in quiet and noise at 500 ms vs 350

ms and 500 ms vs 150 ms showed that the reconstruction accuracy of younger adults is not significantly affected by the integration windows in noise (t[16] = 0.366, p = 0.719 and t[16] = 1.162, p = 0.262 for 500 ms vs 350 ms and 500 ms vs 150 ms, respectively), while in quiet the 500 ms integration window had significantly lower values than 350 ms, but not than 150 ms (t[16] = -3.722, p = 0.002 and t[16] = 0.973, p = 0.345 for 500 ms vs 350 ms and 500 ms vs 150 ms, respectively). Conversely, older adults' ability to track the speech envelope of the foreground is significantly reduced at 150ms in both quiet (t[14] = -0.248, p = 0.807 and t[14] = 3.779, p = 0.002 for 500 ms vs 350 ms and 500 ms vs 150 ms, respectively) and noise (t[14] = 2.064, p = 0.058 and t[14] = 2.512, p = 0.0248 for 500 ms vs 350 ms and 500 ms vs 150 ms, respectively).

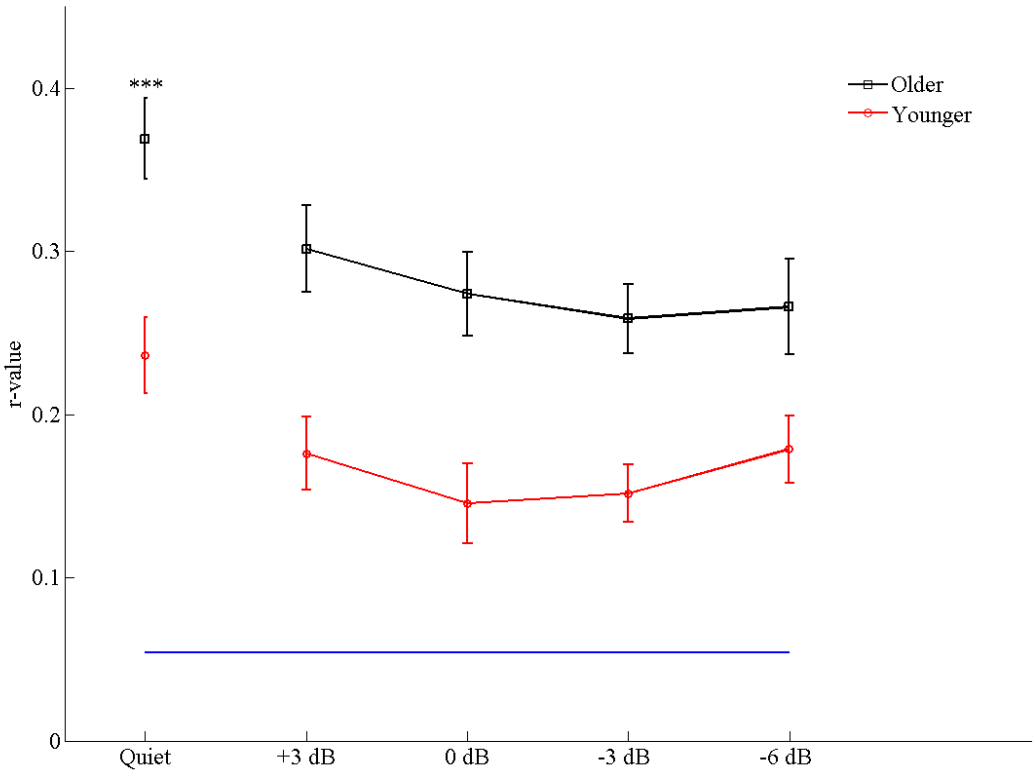

Fig. 24 Reconstrcution accuracy ± 1SE of the speech envelope of the foreground for younger and older adults in quiet and in all the noise conditions tested**.** The blue line shows the noise floor. Older adults' reconstruction accuracy is significantly higher in quiet (p < 0.001). However, as a completing talker is added to the task, the differences between the two age groups are reduced.

Relationships among behavioral, midbrain and cortical data.

Two-tailed Spearman's rank correlation coefficient (ρ) was used to study the correlations among the same measurements as in the first experiment. No significant correlations were found in either younger or older adults in any of the relationships tested.

DISCUSSION

The results of this study provide support for most, but not all, of the initial hypotheses. Behavioral data showed that older adults do have poorer speech understanding in noise than younger adults, despite their normal audiometric hearing thresholds. In midbrain, noise suppresses the response in younger adults to a greater extent than in older adults, while the fidelity of the reconstruction of speech in cortex is higher in older than in younger adults, likely because of an age-related alteration of excitatory/inhibitory mechanisms. Interestingly, even though the speech envelope of the foreground in the noise condition is still better represented in older adults, the addition of a competing talker degrades their responses more than in younger adults. Differently from what was initially hypothesized, no significant associations were found between behavioral and electrophysiological data and between midbrain and cortex.

Midbrain.

Amplitude response.

The greater amplitude decrease in noise in younger adults compared to older adults was unexpected. However, a significant RMS × age group interaction was only significant in the transition region, and may be explained by reduced audibility in the high frequencies

in older adults, given that the transition region is characterized by the presence of a high-frequency burst. These results are consistent with an earlier study that suggested that older adults' high frequency hearing loss might disrupt their ability to encode the high frequency components of a syllable (Presacco et al., 2015). Therefore, the addition of noise would not result in as much additional degradation in the response of older adults as it did in younger adults, because the response in quiet was already degraded.

In younger adults, the loss of amplitude between quiet and noise conditions was also larger in the steady-state region, though no significant RMS × age interaction was observed. The lack of significant differences observed in the steady-state region is consistent with results reported by Parthasarathy et al. (2010), where amplitude modulation following responses (AMFRs) differed in younger and older rats only under specific SNR conditions. Specifically, they observed that at the highest SNR there were no significant differences at any of the modulation frequencies tested, but with a 10 dB loss of SNR the AMFR of younger rats tended to decrease substantially, while older rats' responses showed negligible changes. This is consistent with the results from the second experiment showing significant differences across noise conditions only in younger adults. Additionally, previous studies have shown that hearing loss may lead to an exaggerated representation of the envelope in midbrain (Anderson et al. 2013; Henry et al. 2014). Despite having clinically normal audiometric thresholds up to 4 kHz, most of our older adults have a mild sensorineural hearing loss at higher frequencies (6 and 8 kHz). This mild hearing loss might have potentially contributed to generating an amplitude response big enough to reduce the RMS × age interaction in the steady-state region.

Robustness of the envelope to noise

The correlation analysis did support the initial hypothesis that younger adults' responses should be more robust to noise than those of older adults. Younger adults showed significantly higher correlations when all the noise conditions were grouped together. These differences were significant only in the second experiment, which included more challenging conditions (i.e. -3 and -6 dB SNR). These lower SNRs may better reveal differences between the two age groups, including those due to disruption of periodicity in the encoded speech envelope, which has been suggested to cause a decrease in word identification (Pichora-Fuller et al. 2007).

Cortex

Reconstruction of the speech envelope

The results of the reconstruction of the speech envelope from our second experiment confirm the trend observed in the first experiment, which is that older adults had higher correlation values both in quiet and in noise. An enhanced reconstruction in older adults is consistent with studies showing an exaggerated representation of cortical responses in older adults, both with and without hearing loss. Specifically, Lister et al. (2011), Soros et al. (2009) and Alain et al. (2014) report abnormally higher amplitude for the P1 and N1 peaks in normal hearing older adults compared to normal hearing younger adults, in agreement with results from previous studies that showed that aging might alter inhibitory neural mechanisms in the cortex (de Villers-Sidani et al. 2010; Hughes et al. 2010; Juarez-Salinas et al. 2010; Overton and Recanzone 2016). Interestingly, Chambers et al. (2016) recently showed that recovery from profound cochlear denervation in rats leads to cortical responses

higher than the baseline recorded before inducing auditory neuropathy; this finding reinforces the possibility that auditory neuropathy could play a critical role in the overrepresentation of an auditory stimulus. It is also possible that peripheral hearing loss contributes to problems in the speech encoding process, as several studies have shown that this cortical neural enhancement is exacerbated by hearing loss (Alain et al. 2014; Tremblay et al. 2003). The abovementioned exaggerated cortical response, which can take the form of both better cortical reconstruction and higher peak amplitude (P1 and N1), is perhaps counterintuitive, and in disagreement with the concept of "stronger is better", as observed in the midbrain. However, if we assume that decrease of inhibition leads to larger neural currents, we can hypothesize that this neural enhancement is mainly the result of imbalance between excitatory and inhibitory mechanisms.

Another possible explanation since higher cognitive processes do affect the final representation of speech in the cortex, is related to an inefficient use of cognitive resources and an associated decrease in cortical network connectivity reported in older adults (Peelle et al. 2010). Decreased cortical network connectivity would result in neighboring cortical areas processing the same stimulus independently, instead of collaboratively, which could also lead to over-representation. This would translate to using significantly more energy to accomplish a task that younger adults can complete with much less effort. This explanation would be in agreement with several studies showing that overuse of cognitive resources leads to poorer performance on a secondary task (Anderson Gosselin and Gagne 2011; Tun et al. 2009; Ward et al. 2016).

Importantly, the addition of a competing talker caused a substantial drop of decoding accuracy in older adults, who required a much longer integration time than

younger adults. This finding is consistent with several psychoacoustic (Fitzgibbons and Gordon-Salant 2001; Gordon-Salant et al. 2006) and electrophysiological studies (Alain et al. 2012; Lister et al. 2011), demonstrating that older adults' responses are affected to a greater degree than younger adults when temporal parameters are varied. Specifically, older adults required longer time to process specific temporal acoustic cues, such as voice-onset time, vowel duration, silence duration and transition duration (Gordon-Salant et al. 2006). The degradation of the cortical response from quiet to noise observed in both age groups is also consistent with previous results showing that the evoked response seen in quiet is affected by the presence of noise (Billings et al. 2015). Specifically, response amplitude there, in both younger and older adults, decreases as SNR decreases, consistent with the current findings showing a reduction in reconstruction accuracy within each group. Additionally, this decrease of response in noise could also be related to a different mechanism being used by the auditory system to process information when more than one talker is present. Results from Furman et al. (2013) indeed suggest that hearing in noise relies more on information carried by the low spontaneous rate fibers, whereas in quiet, the information carried by high spontaneous rate neurons is far more important.

Effect of hearing threshold differences and cognitive decline on cortical results.

The possibility that the overrepresentation of the response of older adults in quiet might be due to significant differences in the hearing thresholds cannot be ruled out. In fact, even though the older adults that we tested had clinically normal hearing, all of their thresholds were significantly higher than younger adults ($p < 0.05$), a typical occurrence for the majority of aging studies. Furthermore, several studies have also shown the existence of

age-related cognitive declines (Anderson Gosselin and Gagne 2011; Pichora-Fuller et al. 1995; Surprenant 2007; Tun et al. 2009) that may play a critical role in compromising attentional resources, believed to be critical for a proper representation of the auditory object (Shamma et al. 2011).

Relationships among behavioral, midbrain and cortical data.

The absence of correlations among behavioral and electrophysiological measurements suggest the possibility that our behavioral measurements might not completely account for the presence of temporal processing deficits in the central auditory system. Caution should be used when interpreting the results due to important factors: 1) Behavioral data were collected with four-talker babble as the background noise, while cortical and subcortical data were recorded using a single competing talker. A single competing talker may draw the subjects' attention away from the target to a greater extent than would four-talker babble, given the fact that multiple talkers generate speech without meaning (little informational masking) (Larsby et al. 2008; Tun et al. 2002); 2) Several studies have also shown that the performance in a task varies depending on different features of the masker (i.e. spectral differences, SNR level , etc.) (Calandruccio et al. 2010; Larsby et al. 2008). The speech materials used for the electrophysiological recording were not equated for spectral differences with the speech material used for the speech-in-noise test.

No significant association was found between midbrain and cortical results. Even though previous results showed relationships between weak speech encoding in the midbrain and an overrepresentation of the cortical response (Bidelman et al. 2014), a more recent animal study (Chambers et al. 2016) suggests that the absence of auditory brainstem response does

not necessarily lead to absence of cortical response, suggesting compensatory central gain increases that could help restore the representation of the auditory object at the cortical level. This finding may also explain the lack of association between midbrain and cortex findings in both experiments. It could also be argued that the absence of correlation between midbrain and cortex could be linked to the different stimuli used for the EEG (speech syllable /da/) and MEG (1 minute of speech) task. Additionally, subjects were passively listening to the auditory stimuli in the EEG experiment, while in the MEG subjects were actively engaged in listening to the target speaker.

Concluding remarks.

The results of our studies add compelling evidence to the notion that age-related temporal processing deficits are a key factor in explaining speech comprehension problems experienced by older adults, particularly in noisy environment. Auditory midbrain responses revealed an age-related failure to encode speech syllables in quiet, which translates to an inability to cope with the presence of a background talker. While younger adults adapt to the presence of noise and changes in its loudness, older adults seem to not be affected by different SNRs, suggesting a failure to properly encode both the target and the irrelevant speech. Our study also reveals an overrepresentation of the cortical response, consistent with previous studies (Alain et al. 2014; Lister et al. 2011; Soros et al. 2009); this neural enhancement is reduced with the addition of a competing talker, suggesting that larger cortical responses are not beneficial, and might in fact represent a failure of the brain to properly process speech. Critically, we were unable to find any significant correlations among midbrain and cortex measures. We believe this result brings additional support to

recent findings that suggest the possibility that cortical plasticity can partially restore temporal processing deficits at lower levels of the auditory system (Chambers et al. 2016). This apparent lack of relationship between midbrain and cortex further highlights the relevance of this study, which is the importance of investigating simultaneously different areas of the auditory system to better understand the mechanisms underlying age-related degradation of the speech representation.

CHAPTER IV

Third Study

Objectives

The ability to understand speech is significantly degraded by aging, particularly in noisy environments. One way that older adults cope with this hearing difficulty is through the use of contextual cues. Several behavioral studies have shown that older adults are better at following a conversation when the target speech signal has high contextual content or when the background distractor is not meaningful. Specifically, older adults gain significant benefit in focusing on and understanding speech if the background is spoken by a talker in a language that is not comprehensible to them (i.e. a foreign language). To better understand the neural mechanisms underlying this benefit in older adults, we investigated aging effects on midbrain and cortical encoding of speech when in the presence of a single competing talker speaking in language that is meaningful or meaningless to the listener (i.e., English vs. Dutch). Our results suggest that neural processing is strongly affected by the informational content of noise. Specifically, older listeners' cortical responses to the attended speech signal are less deteriorated when the competing speech signal is an incomprehensible language than when it is their native language. Conversely, temporal processing in the midbrain is affected by different backgrounds only during rapid changes in speech, and only in younger listeners. Additionally, we found a negative correlation between cognitive decline and cortical responses in older adults, suggesting an age-related over (or inefficient) use of cognitive resources that may explain their difficulty in processing speech targets while trying to ignore interfering noise.

INTRODUCTION

Human ability to segregate speech in noisy environments significantly degrades with aging, even when hearing acuity is clinical normal. This difficulty may arise, in part, from different effects of noise on neural speech encoding in younger vs. older adults (Presacco et al. 2016 co-submitted to J. Neurophysiol. along with this manuscript). Specifically, neural synchronization in the midbrain and cortex is deteriorated by noise to a greater extent in older adults. The cortical response also revealed an age-related overrepresentation of response in older adults in both quiet and noise conditions. These results suggest a disruption of the normal balance between excitatory and inhibitory processes and are consistent with several studies showing age-related auditory temporal processing deficits both in the midbrain (Anderson et al. 2012; Caspary et al. 2006; Caspary et al. 1995; Caspary et al. 2005; Parthasarathy and Bartlett 2011; Parthasarathy et al. 2010; Presacco et al. 2015; Walton et al. 1998) and in the cortex (Alain et al. 2014; Getzmann et al. 2016; Getzmann and Naatanen 2015; Lister et al. 2011; Ross et al. 2010; Soros et al. 2009) that would be exacerbated by the presence of hearing loss (Anderson et al. 2013; Henry et al. 2014; Humes and Christopherson 1991; Humes and Roberts 1990; Peelle et al. 2011).

Despite this age-related neural decline, older adults' ability to follow a conversation is not entirely compromised, as they gain additional help from the use of contextual cues. Specifically, several experiments have shown how older adults heavily rely on the context of the conversation to compensate for their speech comprehension problems (Lash et al. 2013; Pichora-Fuller et al. 1995; Rogers et al. 2012; Rogers and Wingfield 2015; Tun et al. 2002). Interestingly, results from Tun et al. (2002) and Brouwer et al. (2012) also suggest that the type of background noise (meaningful vs meaningless) could play a key role in the way older adults process speech. Critically, the results from Tun et al. (2002)'s

study revealed how having a meaningful distractor (comprehensible words spoken in English) impaired the understanding of the target speech to a greater extent than a meaningless distractor (incomprehensible words spoken in Dutch) in older, but not in younger adults. Another important aspect in speech processing is the additional cognitive demand imposed by degraded stimuli that leads older adults to inefficiently use more resources that would otherwise be available for secondary tasks (Anderson Gosselin and Gagne 2011; Tun et al. 2009; Ward et al. 2016).

Altogether, these studies suggest not only that speech-in-noise performance is regulated by a combination of bottom-up and top-down processes that contribute to efficient stream segregation, but also that informational content of noise might have a different impact on the segregation of one auditory stream from another in the cortex, where brain plasticity plays a critical role in building the final representation of the attended sound stream.

This is particularly important in light of a recent study that demonstrated a central compensatory gain mechanism strong enough to restore the representation of sounds at the cortical level in cases of absent auditory brainstem responses induced by auditory neuropathy (Chambers et al. 2016). Critically, attention-related brain plasticity been consistently observed in the cortex in both animal and human experiments (Bidet-Caulet et al. 2007; Choi et al. 2013; Fritz et al. 2003; Lee and Middlebrooks 2011), but less reliably in the midbrain. It is well-known that corticofugal projections from the cortex to midbrain have the ability to regulate and change the activity in lower nuclei (Suga 2008). However, their influence on the kind of short-time plasticity modulated by behavioral tasks, such as attentional tasks, is not well understood. Some recent results bring evidence against the

possibility that responses at such a low auditory level might be controlled by higher cognitive processes (Varghese et al. 2015), while others suggest that there is evidence for the existence of this specific task-related plasticity (Slee and David 2015).

Results from Presacco et al. (2016 co-submitted to J. Neurophysiol. along with this manuscript) demonstrate that both midbrain and cortical responses were degraded in older adults to a greater extent than in younger adults, particularly for the most challenging listening conditions. The current study investigates the differing neural mechanisms underlying age-related deficits in speech-in-noise understanding arising from different types of noise (meaningful vs meaningless), and at different signal-to-noise ratios (SNRs). To minimize the effects of audibility, only clinically normal hearing listeners were included in both the younger and older age groups.

We posit several hypotheses. First, in midbrain responses, we hypothesize that different SNRs, but not different informational content of noise, will significantly affect the fidelity of the response of younger and older adults. Conversely, in the cortex we expect the reconstruction fidelity of the speech envelope to be measurably augmented by the use of meaningless noise in older than in younger adults. Finally, we hypothesize that cognitive decline in older adults will be *negatively* correlated with cortical reconstruction accuracy across subjects. This supposition stems from the observation that the overly high reconstruction fidelity of the speech envelope observed for older adults in our previous study (Presacco et al. 2016 co-submitted to J. Neurophysiol. along with this manuscript) may be a biomarker representing both an imbalance between excitatory and inhibitory processes and an inefficient use of cognitive resource.

MATERIALS AND METHODS

Participants.

Participants comprised 17 younger adults (18 – 27 years, mean ± sd 22.23 ± 2.27, 3 male) and 15 older adults (61 – 73 years old, mean ± sd 65.06 ± 3.30, 5 males) recruited from the Maryland, Washington D.C. and Virginia areas. These participants were the same as those used for the second experiment in our previous study (Presacco et al. 2016 co-submitted to J. Neurophysiol. along with this manuscript), data for which were obtained during the same sessions as this study. All participants had clinically normal hearing (Figure 25) defined as follows: (1) air conduction thresholds ≤ 25 dB HL from 125 to 4000 Hz bilaterally; and (2) no interaural asymmetry (> 15 dB HL difference at no more than two adjacent frequencies). Participants had normal IQ scores [≥ 85 on the Wechsler Abbreviated Scale of Intelligence (Zhu and Garcia 1999)] and were matched on IQ ($F_{(1,30)}$ = 0.660, p = 0.423) and sex (Fisher's exact, $p > 0.05$). In addition, the older adults were screened for dementia on the Montreal Cognitive Assessment (MOCA) (Nasreddine et al. 2005). All participants participated in both the EEG and MEG study, spoke English as their first language and had no understanding of Dutch. Dutch was used because of its similarity to English in terms of phonological inventory and prosodic contours (Collier and Hart 1975). EEG and MEG data for each subject were collected in two different sessions. The Quick Speech-in-Noise test (QuickSIN) (Killion et al. 2004) was used to quantify the ability to understand speech in noise composed of four-talker babble. All procedures were reviewed and approved by the Institutional Review Board (IRB) of the University of Maryland. Participants gave informed consent and were paid for their time.

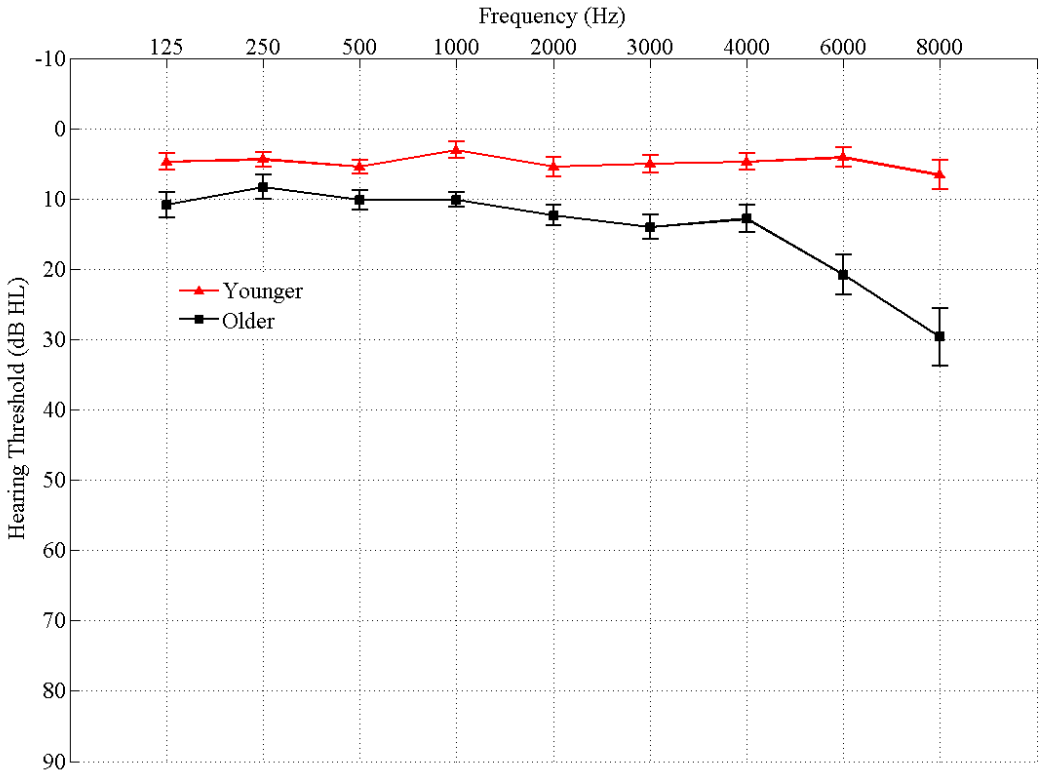

Fig. 25. Audiogram (mean ± 1SE) of the grand averages of younger and older adults. All participants have clinically normal hearing**.**

EEG: Stimuli and recording. A 170-ms /da/ (Anderson et al. 2012) was synthesized at a 20 kHz sampling rate with a Klatt-based synthesizer (Klatt 1980). The stimulus was presented at 75 peak dB SPL diotically with alternating polarities at a rate of 4 Hz through electromagnetically shielded insert earphones (ER-1; Etymotic Research) using Presentation. FFRs were recorded in quiet and in 4 noise conditions: +3 dB, 0 dB, -3 dB and -6 dB SNR, defined as the RMS values between the speech syllable /da/ and the single female competing talker used as the background noise. Each of these four SNR conditions was presented with meaningful (female native English speaker) and meaningless (female native Dutch speaker) background noise at a sampling frequency of 16384 Hz using the Biosemi Active Two acquisition system (Biosemi B.V.) with a standard vertical montage of 5 electrodes (Cz active, forehead ground, earlobe references). During the recording

session (~2 hr), participants sat in a recliner and watched a silent, captioned movie of their choice to facilitate a relaxed yet wakeful state. Two thousand artifact-free sweeps were recorded for each condition from each participant.

Data analysis.

Data recorded with Biosemi were analyzed in MATLAB (MathWorks, version R2011b) after being converted into MATLAB format with the function pop_biosig from EEGLab (Delorme and Makeig 2004). Sweeps with amplitude in the ±30 μV range were retained and averaged in real time and then processed off-line using MATLAB (MathWorks, version R2011b). The time window for each sweep was -47 to 189 ms referenced to the stimulus onset. Responses were digitally bandpass-filtered offline from 70 to 2000 Hz using a 4^{th} order Butterworth filter to minimize the effects of cortical low-frequency oscillations. A final average response was created by averaging the sweeps of both polarities to minimize the influence of cochlear microphonic and stimulus artifact on the response and simultaneously maximize the envelope response (Aiken and Picton 2008; Campbell et al. 2012; Gorga et al. 1985). Root mean square (RMS) values were calculated for the transition (18 ms to 68 ms) and steady-state (68 ms to 170 ms) regions. Correlation (Pearson's linear correlation) between the envelope response in quiet and noise was calculated for each subject to estimate the extent to which noise affects the FFR.

MEG recording. The same subjects recruited for the EEG study participated in the MEG experiment. Participants were asked to attend to one of two stories (foreground) presented diotically while ignoring the other one. The stimuli for the foreground consist of narrated

segments from the book *The Legend of Sleepy Hollow* by Washington Irving. The stimuli for the background consist of segments narrated by an English speaker from the book *A Christmas Carol* by Charles Dickens, and narrated by a Dutch speaker from the book *Aljaska en de Canada-spoorweg* by Anonymous. The foreground was spoken by a male talker, while the background story was spoken by a female talker. Each speech mixture was constructed as described by Ding and Simon (2012) by digitally mixing two speech segments into a single channel, with a duration of 1 minute. Five different conditions, presented to each subject in randomized order, were recorded: quiet, +3dB, 0 dB, -3 dB and -6 dB SNR. The condition in quiet was recorded with two different segments. Three trials of each condition were presented for a total of 30 trials (24 trials in noise and 6 in quiet). As in the case of the EEG part of this study, the 4 conditions in noise were presented in two different scenarios: meaningful (where the competing talker was a female native English speaker narrating the story in English) and meaningless (where the competing talker was a female native Dutch speaker narrating the story in Dutch) noise. The male speaker was always used as the foreground speaker and 8 different segments from the same story were used in order to minimize the possibility that the clarity of the stories could affect the performance of the subjects. In order to maximize the level of attention of the subject on the foreground segment, participants were asked beforehand to count the number of times a specific word or name was mentioned in the story. The sounds (approximately 70 dB SPL when presented with a solo speaker) were delivered to the participants ears with 50 Ω sound tubing (E-A-RTONE 3A; Etymotic Research), attached to E-A-RLINK foam plugs inserted into the ear canal. The entire acoustic delivery system was equalized to give an approximately flat transfer function from 40 to 3000 Hz, thereby encompassing the

range of the presently delivered stimuli. Neuromagnetic signals were recorded using a 157-signal whole head MEG system (Kanazawa Institute of Technology, Kanazawa, Japan) in a magnetically shielded room as described in Ding and Simon (2012).

Finally a set of 50 ms (10 ms cosine ramped), 500 Hz tones was also presented to each participant at the end of each study. Subjects were asked to count the number of occurrences of the tone (100 sweeps). The amplitude of the 100 ms latency peak (M100, or N1m) elicited during this modest task was significantly higher in older adults ($F_{[1,30]} = 5.472$, $p = 0.026$) than younger. Since this result is consistent with previous studies (Alain et al. 2014; Soros et al. 2009), it was not further analyzed.

Data analysis.

Three reference channels were used to measure and cancel the environmental magnetic field by using TS-PCA (De Cheveigné and Simon 2007). MEG data were analyzed off-line using MATLAB. The 157 raw MEG data channel responses were first filtered between 2 - 8 Hz with an order 700 windowed (Hamming) linear-phase FIR filter, then decomposed into N signal components (where $N \leq 157$) using the denoising source separation (DSS) algorithm (de Cheveigne and Simon 2008; Särelä and Valpola 2005). The filtering range of 2 – 8 Hz was chosen based on previous results showing the absence of intertrial coherence above 8 Hz (Ding and Simon 2013) and the importance of the integrity of the modulation spectrum above 1 Hz to understand spoken language (Greenberg and Takayuki 2004). The signal components were extracted for each trial, band-pass filtered between 1 – 8 Hz (Ding and Simon 2012) with a 2^{nd} order Butterworth filter, and averaged over trials. Reconstruction of the envelope was performed using a linear reconstruction matrix

estimated via the Boosting algorithm (David et al. 2007; Ding et al. 2013; Ding and Simon 2013). Success of the reconstruction is measured by the linear correlation between the reconstructed and actual speech envelope. The reconstructed envelope was obtained only from the speech of the single speaker to which the participant was instructed to attend, not of the actual acoustic stimulus. The envelope was computed as the 1 - 8 Hz band-pass filtered magnitude of the analytic signal. In order to optimize the reconstruction fidelity, data were analyzed in a 500 ms integration window (Ding and Simon 2013; Presacco et al. 2016 co-submitted to J. Neurophysiol. along with this manuscript). The noise floor was calculated by using the neural response recorded from each condition to reconstruct the speech envelope of a different stimulus than was used during this response.

Cognitive test. The Flanker Inhibitory Control and Attention Test of the National Institutes of Health Cognition Toolbox was used to measure executive function (ability to inhibit visual attention to irrelevant tasks). Participants were shown a series of arrows and were asked to determine as quickly as possible the direction of the middle arrow by pressing the space bar. The unadjusted scale score was used to compare age-related differences.

Statistical analyses.

All statistical analyses were conducted in SPSS version 21.0 (SPSS). Split plot ANOVAs were used to test for age group × condition interactions for the RMS values of the FFR response in the time domain and for correlation values calculated for the cortical data. The Greenhouse-Geisser test was used when the Mauchly's sphericity test was violated. A paired t-test was used for within-subject group analysis for the correlation values and

amplitudes for the cortical data, while one-way analyses of variance (ANOVAs) were used to analyze the RMS values of the amplitude of the FFR and the correlation values and the correlation for the cortical data. The non-parametric Mann-Whitney U test was used in place of one-way ANOVA when Levene's test of Equality of Variances was violated. A one-sample t-test was used to evaluate the slopes of the RMS calculated for the four noise conditions in the transition region. Two-tailed Spearman's rank correlation (ρ) was used to evaluate the relationships among cognitive score and midbrain and cortical parameters. Fisher's z transformation was applied to all the correlation values calculated for the midbrain and cortical analysis. The false discovery rate (FDR) procedure (Benjamini and Hochberg 1995) was applied to control for multiple comparisons where appropriate.

RESULTS

Speech Intelligibility.

Younger adults (mean ± std = -0.573 ± 1.13) scored significantly better ($F_{[1,30]} = 10.613$, $p = 0.003$) than older adults (mean ± std = 0.8 ± 1.25), suggesting that older adults' performance in noise may decline compared to younger adults even when audiometric thresholds are clinically normal.

Midbrain (EEG).

Amplitude analysis.

Figure 26 displays the RMS values of younger and older adults for each SNR condition tested with meaningful and meaningless noise.

Transition region. A paired t-test showed significantly higher RMS values with meaningless noise in both younger ($t_{[16]} = -2.911$, $p = 0.01$, $t_{[16]} = -2.234$, $p = 0.04$, $t_{[16]} = -$

2.522, $p = 0.02$ and $t_{[16]} = -4.283$, $p < 0.001$ for +3 dB, 0 dB, -3 dB and -6 dB respectively) and in older adults only at 0 dB ($t_{[14]} = -4.514$, $p < 0.001$). A regression analysis was also carried out to fit the conditions in noise. A one-sample t-test showed that the slopes for both meaningful and meaningless noise are significantly different from zero in younger adults ($t_{[16]} = 4.763$, $p < 0.001$ and $t_{[16]} = 4.247$, $p < 0.001$ for meaningful and meaningless noise, respectively), while in older adults only the slopes of the meaningless noise were significantly different from zero ($t_{[14]} = 0.886$, $p = 0.391$ and $t_{[16]} = 3.910$, $p = 0.002$ for meaningful and meaningless noise, respectively). Additionally, the regression analysis also revealed a ~4.6 dB neural advantage for meaningless noise over meaningful noise (blue line in the inset of Figure 2). Repeated measures ANOVA showed no noise type × age interaction in either the conditions tested (all p values > corrected significance threshold). Repeated measures ANOVA showed a condition × age interaction between quiet and noise at -3 dB ($F_{[1,30]} = 6.264$, $p = 0.018$) and -6 dB ($F_{[1,30]} = 6.696$, $p = 0.015$), but not at the other conditions tested ($F_{[1,30]} = 1.125$, $p = 0.297$ and $F_{[1,30]} = 0.333$, $p = 0.568$, and for +3 dB and 0 dB respectively),with meaningful noise. Conversely, with meaningless noise the interaction was found at -3 dB ($F_{[1,30]} = 8.097$, $p = 0.008$), but not at the average of +3 dB ($F_{[1,30]} = 1.294$, $p = 0.264$), 0 dB ($F_{[1,30]} = 1.986$, $p = 0.169$) and -6 dB ($F_{[1,30]} = 1.784$, $p = 0.192$). A one-way ANCOVA (using the condition in quiet as covariate, as younger adults have significantly higher RMS values $F_{[1,30]} = 4.255$, $p = 0.048$) showed no significant differences between younger and older adults in any of the conditions tested with meaningful ($F_{[1,29]} = 0.007$, $p = 0.936$, $F_{[1,29]} = 0.296$, $p = 0.590$, $F_{[1,29]} = 1.941$, $p = 0.174$ and $F_{[1,29]} = 2.511$, $p = 0.124$ for +3 dB, 0 dB, -3 dB and -6 dB respectively) and

meaningless ($F_{[1,29]} = 0.195$, $p = 0.662$, $F_{[1,29]} = 0.278$, $p = 0.602$, $F_{[1,29]} = 3.779$, $p = 0.062$ and $F_{[1,29]} = 0.077$, $p = 0.783$ for +3 dB, 0 dB, -3 dB and -6 dB respectively) noise.

Steady-state region. A paired t-test showed no significance differences between meaningful and meaningless noise in either age groups (all p values > corrected significance threshold). Repeated measures ANOVA also show no condition × age interaction between quiet and the two types of noise tested (all p values > corrected significance threshold). Given these results, the RMS for meaningful and meaningless noise were collapsed together in one single analysis. A one-way ANOVA showed significantly higher RMS value in younger adults than older adults in all the conditions tested ($F_{[1,62]} = 11.632$, $p = 0.001$, $F_{[1,62]} = 16.606$, $p < 0.000$, $F_{[1,62]} = 9.813$, $p = 0.003$ and $F_{[1,62]} = 14.840$, $p < 0.000$ for +3 dB, 0 dB, -3 dB and -6 dB respectively).

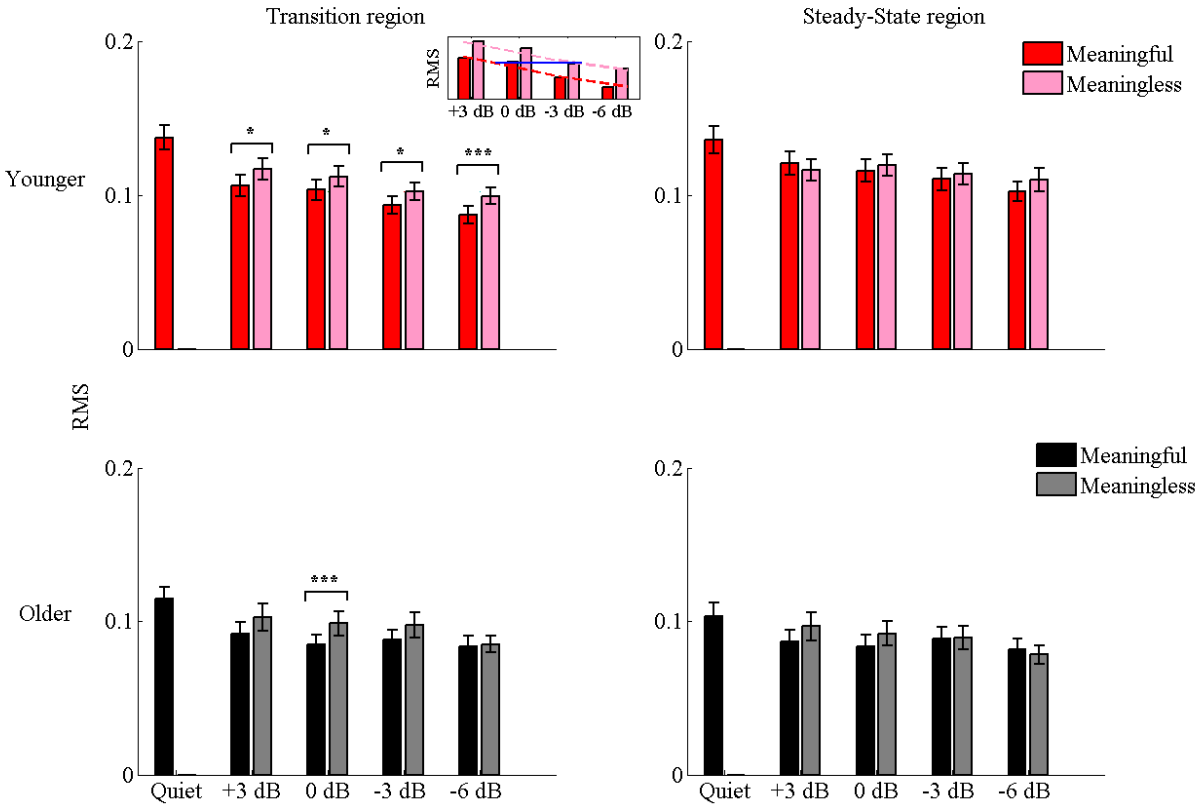

Fig. 26: RMS values ± 1SE for younger (top) and older (bottom) adults in the transition (left) and steady-state (right) regions for all the conditions tested. Different informational content of noise affects only the

transition response and mainly in younger adults. The blue line in the inset shows the neural advantage (~4.6 dB) for meaningless noise. A significant effect of the different informational content of noise was seen only in the transition region in all the conditions tested in younger adults ($p < 0.05$ at +3 dB, 0 dB and -3 dB and $p < 0.001$ at -6 dB), and only at 0 dB in older adults ($p < 0.001$). Significantly higher RMS values in younger adults than older adults in all the conditions tested were found only in the steady-state region. *$p < 0.05$, ***$p < 0.001$

Correlation analysis. In order to analyze the robustness of the response in noise, we linearly correlated (Pearson correlation) the average response (Figure 27) obtained in quiet with the ones obtained in noise, for both the transition and steady-state regions for each subject. Repeated measures ANOVA showed no noise type × age interaction (all p values > 0.05) between meaningful and meaningless noise in all the conditions tested in either regions. A paired t-test also showed no significant differences between meaningful and meaningless noise in either of the conditions tested (all p values > corrected significance threshold) in either region. Given these results, the correlation values for meaningful and meaningless noise were collapsed together for further analyses for both the transition and the steady-state region. The Mann-Whitney U test was used to study the differences between groups. No significant differences were found in the transition region in all the conditions tested (U [62] = 398, Z = -1.507, p = 0.132, U [62] = 424, Z = -1.157, p = 0.247, U [62] = 439, Z = -0.955, p = 0.339 and U [62] = 442, Z = -1.184, p = 0.236 for +3 dB, 0 dB, -3 dB and -6 dB respectively). Conversely, in the steady-state region significantly higher r-values were found in younger adults at all the conditions tested (U [62] = 314, Z = -2.637, p = 0.008, U [62] = 333, Z = -2.381, p = 0.017, U [62] = 349, Z = -2.166, p = 0.03 and U [62] = 329, Z = -2.435, p = 0.015 for +3 dB, 0 dB, -3 dB and -6 dB respectively).

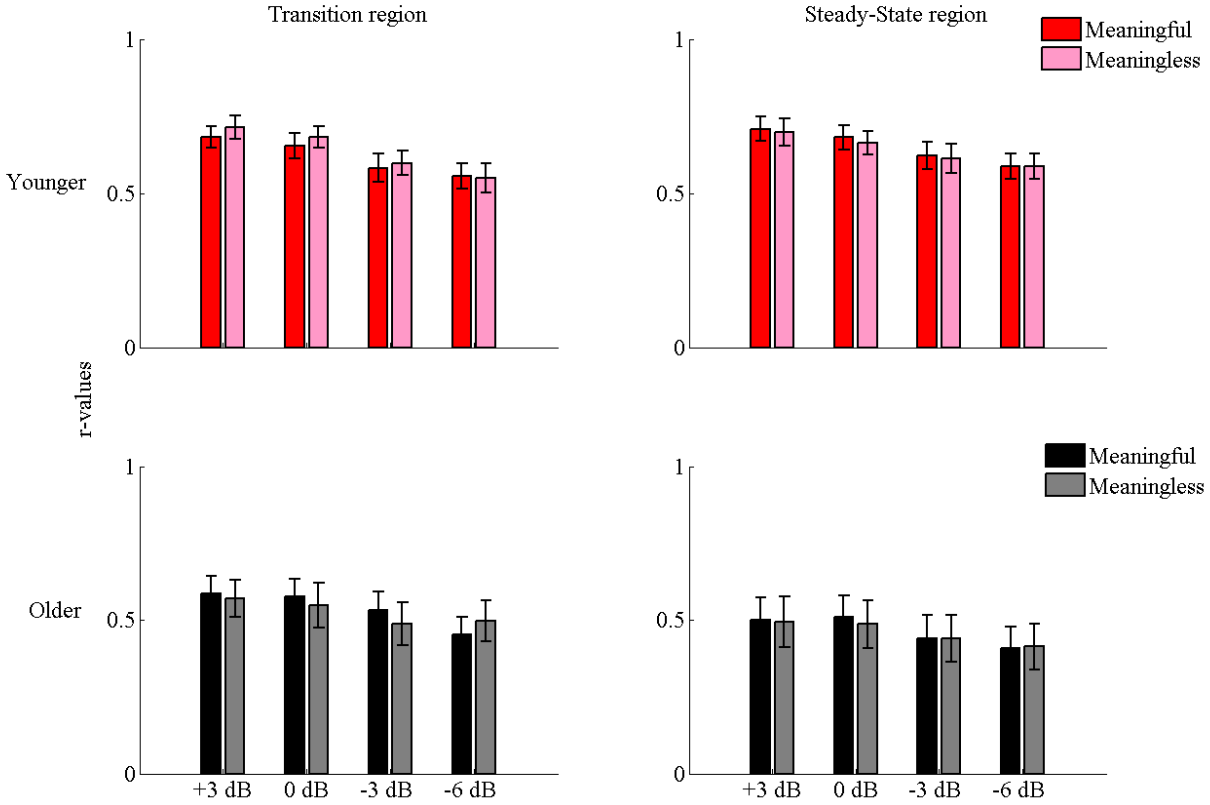

Fig. 27: r-values ± 1SE for younger (top) and older (bottom) adults in the transition (left) and steady-state (right) regions for all the conditions tested. Results showed no significant effect of the type of noise in either younger or older adults for all the conditions tested. Significantly higher r-values were found in younger adults at all the conditions tested.

Cortex (MEG): Reconstruction of the speech envelope.

The ability to reconstruct the low-frequency speech envelope from cortical activity is a measure of the fidelity of the neural representation of that speech envelope (Ding and Simon 2012). Figure 28 displays the grand average ± standard error of the reconstruction accuracy for younger and older adults for all the conditions tested with meaningful and meaningless noise. Data recorded in quiet were from two different 1-minute excerpts. Since no significant interactions ($F_{[1,30]} = 2.340$, $p = 0.137$) and differences between the two excerpts in quiet in both younger ($t_{[16]} = -0.078$, $p = 0.939$) and older ($t_{[14]} = 1.776$, $p = 0.098$) adults were found, their r-values were averaged together for further analyses. A one-

way ANOVA showed significantly higher correlation values in quiet in older adults ($F_{[1,30]}$ = 10.705, p = 0.003). A t-test showed significant differences between meaningful and meaningless noise at -6 dB only in older adults ($t_{[14]}$ = -3.659, p = 0.003). Given these results, a type of noise × age interaction was investigated at -6 dB where repeated measures ANCOVA (with correlation in quiet used a covariate) showed a significant interaction ($F_{[1,29]}$ = 7.008, p = 0.013). A paired t-test showed that reconstruction fidelity was significantly higher than the noise floor in both younger and older adults at all the conditions tested (all p values < 0.01). A follow-up one-way ANCOVA (with correlation in quiet used a covariate) showed no significant differences between younger and older adults at any of the noise conditions tested with meaningful ($F_{[1,29]}$ = 3.136, p = 0.087, $F_{[1,29]}$ = 3.187, p = 0.085, $F_{[1,29]}$ = 4.782 p = 0.037, $F_{[1,29]}$ = 0.020, p = 0.890 for +3 dB, 0 dB, -3 dB and -6 dB respectively) and meaningless ($F_{[1,29]}$ = 1.192, p = 0.284, $F_{[1,29]}$ = 4.061, p = 0.053, $F_{[1,29]}$ = 3.491 p = 0.072 and $F_{[1,29]}$ = 6.058, p = 0.020 for +3 dB, 0 dB, -3 dB and -6 dB respectively) noise. Despite a larger drop in reconstruction fidelity in older adults, their correlations values were still better than younger adults at all the conditions tested (all p < 0.05 with one-way ANOVA applied without covariate).

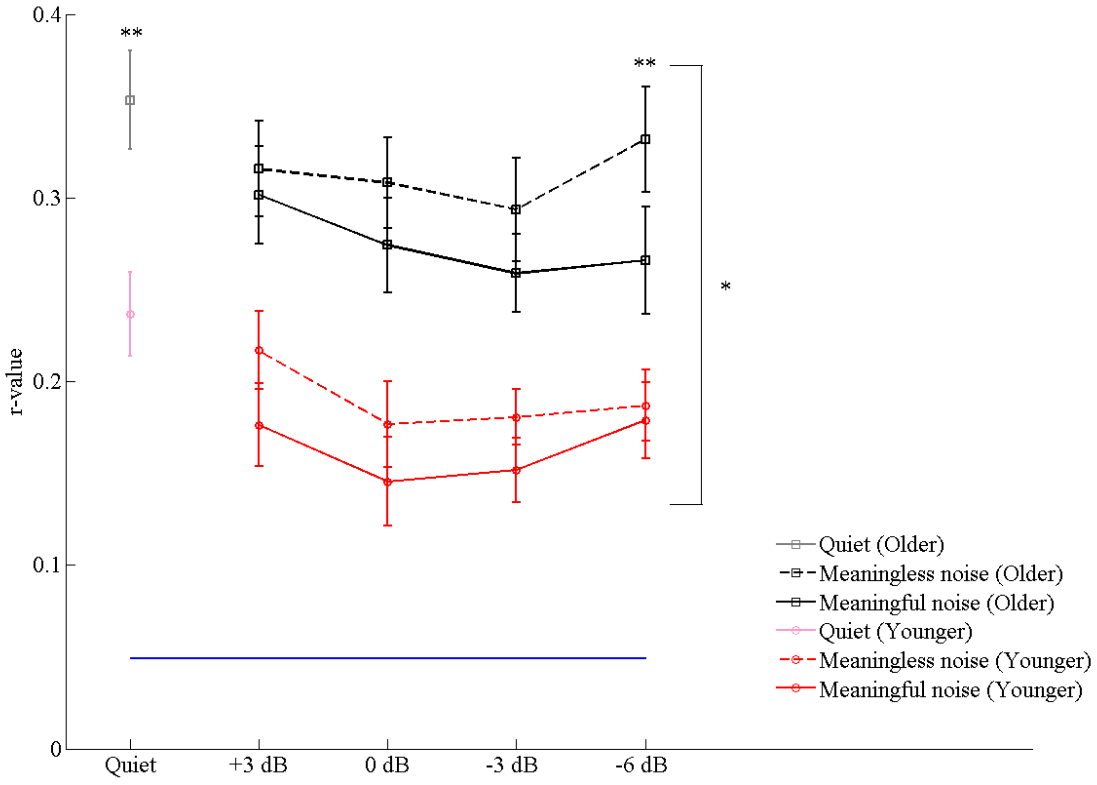

Fig. 28 Reconstruction accuracy ± 1SE of the speech envelope of the foreground for younger and older adults in quiet and in meaningful and meaningless noise**.** The blue line shows the noise floor. Older adults' reconstruction fidelity is better than youngers' at all the conditions tested. Only older adults show significant improvemnt at -6 dB when competing talker was a Dutch speaker. *$p < 0.05$, **$p < 0.01$

Relationships among cognitive test, midbrain and cortical data.

The Flanker Inhibitory control test showed significantly higher scores for younger adults than for older adults ($F_{[1,30]} = 27.375$, $p < 0.001$). This score was used to study the correlation (two-tailed Spearman's rank correlation (ρ)) with the average correlation across all the noise conditions between quiet and noise in the steady-state region of midbrain responses and with average cortical decoding accuracy across all the noise conditions. Significant negative correlations (lower score associated with higher reconstruction accuracy) were found between the Flanker Inhibitory control test score and the cortical response (ρ = -0.621, p = 0.013 and ρ = -0.639, p = 0.01 for meaningful and meaningless

noise respectively) in older, but not in younger adults (ρ = 0.431, p = 0.084 and ρ = 0.348, p = 0.171 for meaningful and meaningless noise respectively). No significant correlation was found between the Flanker Inhibitory control test score and midbrain response in either younger or older adults (all $p > 0.05$). Similarly, no significant correlations were found between midbrain and cortical responses in either younger or older adults ($p > 0.05$). The scatter plot in Figure 29 shows the data for each subject.

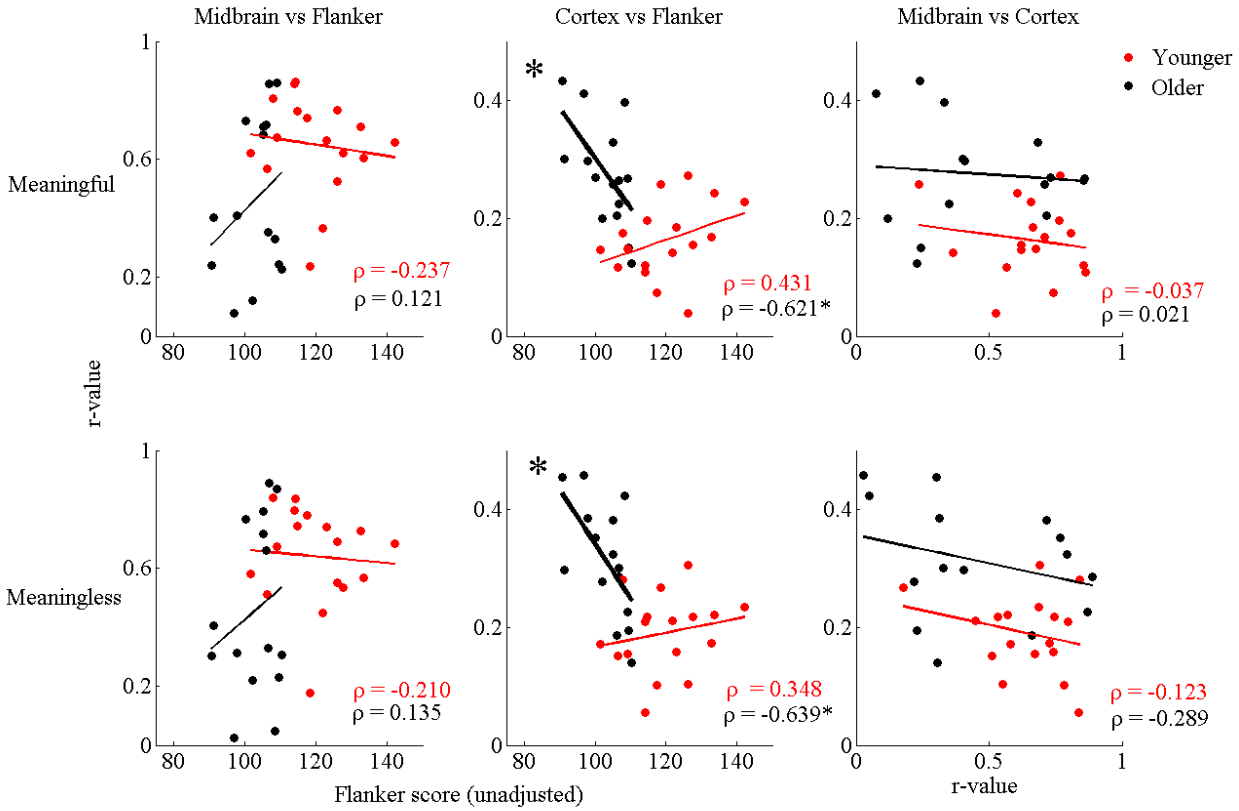

Fig. 29 Scatter plots of the (cognitive) Flanker score, midbrain and cortex r-values for each single subject tested in meaningful and meaningless noise. Significant negative correlations were found between Flanker score and cortical r-values only in older adults (center column, black symbols only).

Discussion.

The results of this study provide support for most, but not all, of our initial hypotheses. Behavioral data showed that older adults do have poorer speech understanding in noise than younger adults, despite their normal audiometric hearing thresholds. In midbrain,

younger adults' responses are less affected by meaningless than meaningful noise in the transition region, an effect not seen in older adults. Consistent with the second experiment reported in our previous study (Presacco et al. 2016 co-submitted to J. Neurophysiol. along with this manuscript), the fidelity of the reconstruction of speech in cortex remains higher in older than in younger adults even with meaningless noise. As predicted, the use of meaningless noise significantly benefited only older adults, as seen most readily in the most challenging conditions. Cortical reconstruction results were also significantly correlated with cognitive scores in older adults as hypothesized, in that the higher reconstruction accuracy the lower their cognitive score.

Midbrain.

Amplitude response.

Transition region. Contrary to what might have been expected, younger adults showed significant differences in the midbrain for all noise conditions tested when meaningless vs. meaningful noise was played in the background, suggesting a substantial effect of top-down mechanisms in the midbrain. The level of influence of higher level cognitive processes on midbrain has led to varying results in the literature, some showing strong effect of attention (Slee and David 2015), with others failing to find any attentional-related change (Varghese et al. 2015). In our study, participants were asked to passively listen to auditory stimuli while watching a silent movie, which should minimize any potential effects of attention. However, it is possible that the midbrain could have encoded FFR in different ways due to the different natures of the background noise. Interestingly, the neural advantage estimated for meaningless noise over meaningless noise was ~4.6 dB across all

the noise conditions tested. The same level of effect was not observed in older adults (only significant at for 0 dB), possibly due to their problems in encoding the high frequency burst, as reflected by significant RMS × age group interactions found in the two most challenging conditions in the transition region, consistent with our previous studies (Presacco et al. 2015; Presacco et al. 2016 co-submitted to J. Neurophysiol. along with this manuscript). Critically, older adults' slope of the line that best fits the RMS conditions of meaningful noise is not significantly different from zero, suggesting that their midbrain response in this condition is not dependent on SNR. This lowered response (compared to younger listeners), and the weak differences in response between the noise conditions, contribute additional evidence of temporal deficits in the transition region.

Steady-state region. In contrast, no noise type effect was found in the steady-state region for either age group.

Robustness of the envelope to noise

Differently from what observed with the amplitude analysis, no significant differences in quiet-to-noise correlations were found for noise backgrounds of different informational content in younger adults in the transition region. The correlation analysis also supported the initial hypothesis that younger adults' responses should be more robust to noise than those of older adults and that the type of noise would have no effect on the response consistency. Younger adults showed significantly higher correlations at all the noise conditions tested in the steady-state region only, reinforcing the existence of a disruption of periodicity in the encoded speech envelope in older adults (Anderson et al. 2012; Mamo

et al. 2016; Pichora-Fuller et al. 2007; Presacco et al. 2015; Presacco et al. 2016 co-submitted to J. Neurophysiol. along with this manuscript).

Cortex

Reconstruction of the speech envelope

A critical part of our experiment was to investigate how different informational content of noise affects the ability to reconstruct the speech envelope of the attended speaker. As hypothesized, reconstruction accuracy was higher for both age groups in the presence of meaningless noise compared to meaningful noise at all SNRs; however, as the SNR decreased, older adults relied more on the type of background than younger adults to process speech, as revealed by a significant correlation × age interaction at -6 dB. These observations are in agreement with a previous study in which older and younger adults were challenged to recall target words in the presence of a meaningful (English) and meaningless (Dutch) distractor (Tun et al. 2002). Consistent with our results, while younger adults' performance did not significantly differ between the two noise conditions, older adults' performance was significantly enhanced when meaningless noise was used as distractor. The results of the reconstruction of the speech envelope also showed an enhanced reconstruction in older adults, which is consistent with studies showing an exaggerated representation of cortical responses in older adults, both with and without hearing loss (Alain et al. 2014; Lister et al. 2011; Presacco et al. 2016 co-submitted to J. Neurophysiol. along with this manuscript; Soros et al. 2009; Tremblay et al. 2003). As discussed by Presacco et al. (2016 co-submitted to J. Neurophysiol. along with this manuscript), this overrepresentation of the response to speech (even in quiet), may result

from a processing deficit or imbalance between excitatory and inhibitory mechanisms. Cognitive resource use may also play a role. Peelle et al. (2010) argue that aging specifically affects the efficient use of cognitive resources because of decreased cortical network connectivity. This in turn would cause neighboring cortical areas to process the same stimulus independently, instead of collaboratively, which could also lead to over-representation. Furthermore, several studies have also suggested that aging might alter the balance between inhibitory and excitatory neural mechanisms in the cortex (de Villers-Sidani et al. 2010; Hughes et al. 2010; Juarez-Salinas et al. 2010; Overton and Recanzone 2016), which in turn might lead to a stronger cortical response. The addition of a competing talker caused a substantial drop of decoding accuracy in older adults to the point that significant differences were no longer seen between the two age groups. This is consistent with recent results (Getzmann et al. 2016; Getzmann and Naatanen 2015) showing age-related changes in event-related potentials recorded in a simulated "cocktail party" scenario that could help explain the difficulties experienced by older adults in effectively segregating and encoding speech streams in noise conditions.

Relationships among cognitive, midbrain and cortical data.

In our prior study (Presacco et al. 2016 co-submitted to J. Neurophysiol. along with this manuscript), one of our open questions was the role of previously reported age-related cognitive decline (Anderson Gosselin and Gagne 2011; Pichora-Fuller et al. 1995; Surprenant 2007; Tun et al. 2009) in explaining the overrepresentation of the cortical response. Here we address this issue by analyzing the correlation between cognitive task and cortical response. Results from this analysis showed that older adults' cognitive decline

in executive function is negatively correlated with speech envelope reconstruction. This negative correlation supports our hypothesis that higher reconstruction accuracy does not translate into better speech understanding, but it is more likely the result of increased neural currents caused by an imbalance between excitatory and inhibitory mechanisms. The failure of these mechanisms might explain the need for older adults to use supplementary cognitive resources to complete the task, characterized by activation of larger areas of the brains, including the prefrontal cortex (Wong et al. 2010) and the cingulo-opercular network (Vaden et al. 2015). However, because of their abovementioned age-related cognitive decline, the level of resources available to them is significantly depleted, making a dual task (e.g. focusing on the target speech while ignoring the background talker) hard to carry out. Additionally, reduced coherence among brain regions involved in speech comprehension might also lead to inefficient use of these cognitive resources, making the speech-in-noise task even more challenging to accomplish (Peelle et al. 2010).

The absence of correlation between midbrain and cortical measurement gives additional support to the hypothesis that compensatory central gain increases help restore the representation of an auditory stimulus at the cortical level, even in the absence of an auditory brainstem response (Chambers et al. 2016). These results would also suggest that this central gain mechanism is restricted in what it can accomplish, e.g., recovery of spike rate encoding, but not the encoding of precise spike timing. An alternative explanation, however, is that the different natures of the task and stimuli used to elicit FFR and cortical responses contribute to this lack of correlation (Presacco et al. 2016 co-submitted to J. Neurophysiol. along with this manuscript)

Concluding remarks.

The overall results of our study gives compelling support to our hypotheses of the existence of an age-related effect for different informational content of noise on the auditory response. These findings are consistent with behavioral studies showing that older adults make use of contextual cues to a greater extent than younger adults to compensate for their problems in understanding speech, particularly in noisy conditions. The presence of meaningless noise led to significantly enhanced representation of the cortical response only in older adults. The strong correlation between cognitive decline and overrepresentation of the cortical response in older adults reinforces our hypothesis that larger cortical responses are not advantageous and might represent an indication of a failure of inhibitory processes and of an inefficient use of cognitive resources.

CHAPTER V

Extension of the third study to older adults with hearing loss (OHI)

The experimental protocol of the third study was used to collect data from a small subset of hearing impaired older adults (OHI) in the attempt to generate preliminary results that could be used for a future experiment that would involve participants with some degree of hearing loss. This investigation is particularly important to understand the extent to which speech in noise problems are affected by peripheral hearing loss and temporal auditory processing deficits in the central auditory system. Results of this analysis are compared with normal hearing younger (YNH) and older (ONH) adults that participated in the third study.

Preliminary results from a small subset

RESULTS

Audiogram and Speech Intelligibility.

Figure 30 shows the audiogram and the speech intelligibility scores for YNH, ONY and OHI. Peripheral hearing loss causes speech-in-noise performance to drastically decline (mean ± std = 3.9 ± 1.79 dB).

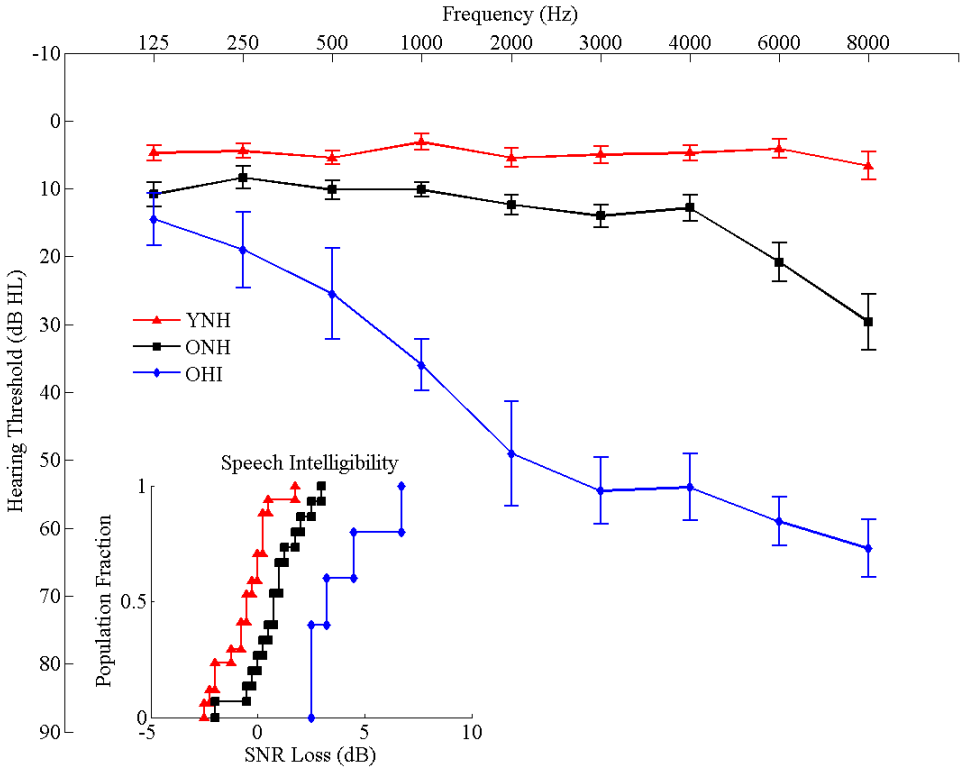

Fig. 30 Audiogram (mean ± 1SE) of the grand averages of YNH, ONH and OHI.

Midbrain (EEG).

Amplitude analysis.

Figure 31 shows the RMS values ± 1SE for YNH, ONH and OHI. Similar RMS values were found between ONH and OHI in the transition region. However, in the steady-state region OHI seems to generate higher amplitude responses than ONH when meaningful noise was used (0.095 vs 0.086 for +3 dB, 0.101 vs 0.083 in 0 dB, 0.1 vs 0.088 in -3 dB and 0.086 vs 0.081 in -6 dB). Similar values were found in quiet (0.102 vs 0.103). Despite the small number of OHI subjects tested, this trend seems to be consistent with previous results showing an overrepresentation of the midbrain response with individuals with hearing loss (Anderson et al. 2013; Henry et al. 2014).

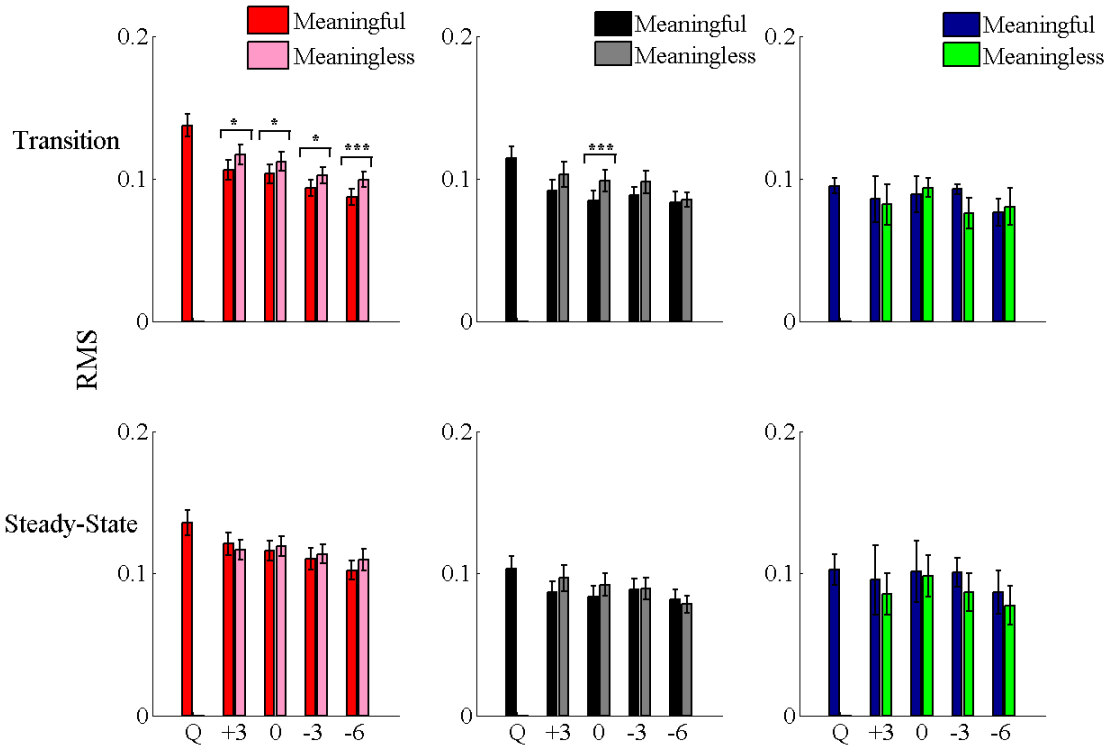

Fig. 31 RMS values ± 1SE for YNH (left), ONH (middle) and OHI (right) adults in the transition (top) and steady-state (bottom) regions for all the conditions tested. Preliminary results show that OHI have higher RMS values than ONH in the steady-state region when meaningful noise was used.

Correlation analysis.

Figure 32 shows the r-values values ± 1SE for YNH, ONH and OHI. Differently from what observed with the RMS analysis, no differences were seen between ONH and OHI in either the regions or conditions tested.

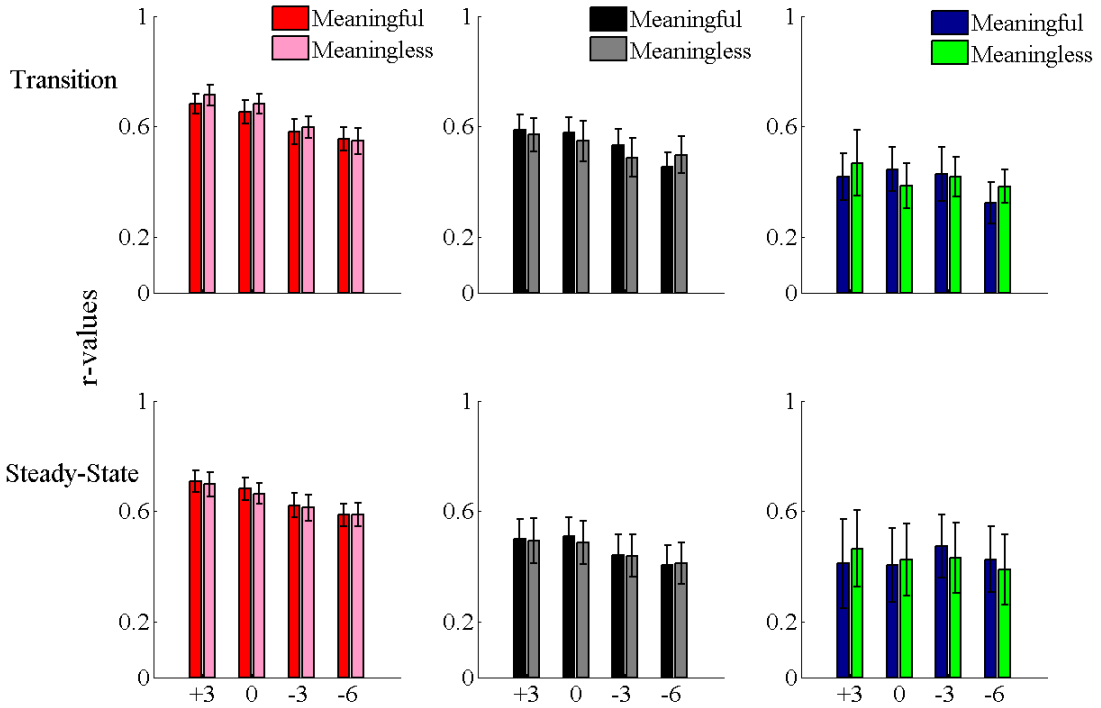

Fig. 32 r-values ± 1SE for YNH (left), ONH (middle) and OHI (right) adults in the transition (top) and steady-state (bottom) regions for all the conditions tested. Preliminary results show similar r-values in ONH and OHI.

Cortex (MEG).

Figure 33 shows the r-values values ± 1SE for YNH, ONH. The standard error for OHI was not plotted because its high value, due to the small sample size; this would have obfuscated the results. An overrepresentation is seen in both ONH and OHI, suggesting that this phenomenon is more likely to be related to an imbalance between excitatory and inhibitory mechanisms and a possible inefficient use of cognitive resources, rather than to peripheral hearing loss. In fact, if hearing loss were the primary cause of this exaggerated response, we would expect to see in ONH correlation values comparable to, or lower than,

YNH. A drop in hearing sensitivity seems to exacerbate, rather than triggering, the age-related phenomena observed in ONH.

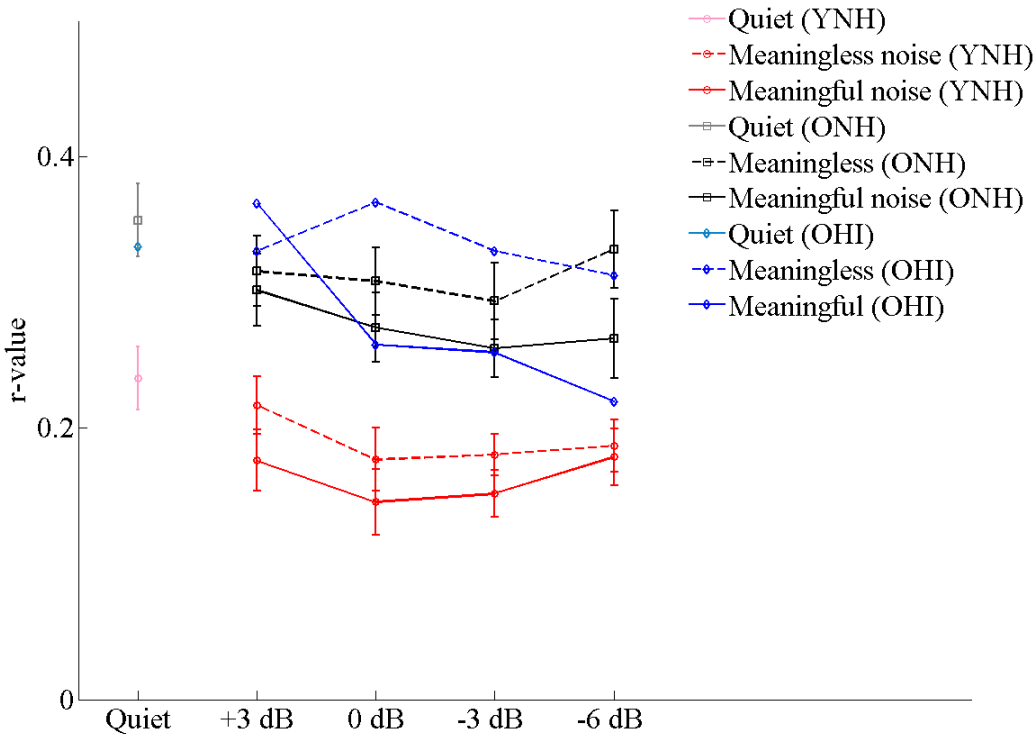

Fig. 33 Reconstrcution accuracy ± 1SE of the speech envelope of the foreground for YNH, ONH and OHI in quiet and in meaningful and meaningless noise**.** The overrepresentation of the cortical response is seen in both ONH and OHI.

Summary of the experiments and future directions

The goals of this dissertation were to understand how age affects neural auditory mechanisms and at which level in the auditory system these changes were particularly relevant for explaining speech-in-noise problems. Our results bring additional evidence to the hypothesis that central auditory processes play a key role in explaining speech recognition problems that older adults experience on a daily basis, particularly in noisy environments. Temporal processing deficits are evident in the midbrain, where the fine temporal precision required for the neurons to fire synchronously is disrupted, leading to lower amplitude responses and worse consistency between the responses in quiet and noise. Despite this degradation at the lower level of the auditory pathway, the cortex does an exceptional job of compensating for these problems in the midbrain and to create a fairly accurate representation of the auditory objects in favorable conditions (i.e. quiet). The presence of a competing talker adds an additional layer of complexity in the analysis that seems to be particularly deleterious in extremely challenging conditions (i.e. -6 dB). Interestingly, older adults use the context (meaningful vs meaningless noise) as a way to cope with their speech-in-noise problems. Specifically, we found that a foreign language is better ignored by older adults and leads to a better representation of the speech envelope. Our results also suggest that the belief “the bigger the better” should be revisited when dealing with older adults. Their cortical response is overrepresented if compared with our normal hearing younger adults. We found that abnormally high reconstruction fidelity of the speech envelope is strongly correlated with poor scores in cognitive tasks. This result would be in agreement with the general knowledge that older adults are unable to efficiently use their cognitive resources, possibly due to degraded neural connections

between different areas of the brain. This cognitive decline along with the presence of the imbalance between excitatory and inhibitory mechanisms may be the main cause of the overrepresentation of the cortical response. An additional goal of our study was to investigate a possible interaction between processing carried out in the midbrain and cortex. Our findings revealed no correlation between midbrain and cortex in either age group. This result is perhaps not surprising, given a recent study showing the existence of a compensatory central gain increase strong enough to help restore the representation of the auditory object at the cortical level, even in the absence of a brainstem response.

Despite the strong evidence that our studies bring in support of the failure of central auditory mechanisms as a possible main reason to explain the age-related degradation of auditory responses, it is clear that peripheral hearing loss also contributes to poor performance in speech-in-noise tasks. In order to address this problem, we started collecting data from older adults with hearing loss in an attempt to dissociate the effects of poor peripheral hearing from the effect of central auditory deficits. Our preliminary results show a possible influence of peripheral hearing loss in the midbrain, but not so much in the cortex. A bigger sample size will be needed to draw solid conclusions.

Future directions will be focused on using the biomarkers identified in this study to assess the efficacy of auditory training techniques. It is evident that hearing sensitivity cannot be restored, if not with the use of hearing aid technology. However, it is possible that training might help older adults partially restore their temporal auditory deficits and thus improve their understanding of speech in noise. A joint collaboration between several hearing research faculties at the University of Maryland is currently collecting preliminary data in attempt to devise experiments aimed at assessing the efficacy of auditory training.

REFERENCES

Aiken SJ, and Picton TW. Envelope and spectral frequency-following responses to vowel sounds. *Hearing research* 245: 35-47, 2008.

Aiken SJ, and Picton TW. Envelope Following Responses to Natural Vowels. *Audiology and Neurotology* 11: 213-232, 2006.

Alain C, McDonald K, and Van Roon P. Effects of age and background noise on processing a mistuned harmonic in an otherwise periodic complex sound. *Hearing research* 283: 126-135, 2012.

Alain C, Roye A, and Salloum C. Effects of age-related hearing loss and background noise on neuromagnetic activity from auditory cortex. *Front Syst Neurosci* 8: 8, 2014.

Altschuler RA, Dolan DF, Halsey K, Kanicki A, Deng N, Martin C, Eberle J, Kohrman DC, Miller RA, and Schacht J. Age-related changes in auditory nerve–inner hair cell connections, hair cell numbers, auditory brain stem response and gap detection in UM-HET4 mice. *Neuroscience* 292: 22-33, 2015.

Anderson Gosselin P, and Gagne JP. Older adults expend more listening effort than young adults recognizing speech in noise. *Journal of speech, language, and hearing research : JSLHR* 54: 944-958, 2011.

Anderson S, Parbery-Clark A, White-Schwoch T, Drehobl S, and Kraus N. Effects of hearing loss on the subcortical representation of speech cues. *The Journal of the Acoustical Society of America* 133: 3030-3038, 2013.

Anderson S, Parbery-Clark A, White-Schwoch T, and Kraus N. Aging affects neural precision of speech encoding. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 32: 14156-14164, 2012.

Bear MF, Connors BW, and Paradiso MA. *Neuroscience: Exploring the Brain*. Baltimore: Lippincott Williams & Wilkins, 2006.

Bidelman GM, and Krishnan A. Effects of reverberation on brainstem representation of speech in musicians and non-musicians. *Brain research* 1355: 112-125, 2010.

Bidelman GM, Villafuerte JW, Moreno S, and Alain C. Age-related changes in the subcortical-cortical encoding and categorical perception of speech. *Neurobiology of aging* 35: 2526-2540, 2014a.

Bidelman GM, Weiss MW, Moreno S, and Alain C. Coordinated plasticity in brainstem and auditory cortex contributes to enhanced categorical speech perception in musicians. *The European journal of neuroscience* 2014b.

Bidet-Caulet A, Fischer C, Bauchet F, Aguera P-E, and Bertrand O. Neural Substrate of Concurrent Sound Perception: Direct Electrophysiological Recordings from Human Auditory Cortex. *Frontiers in human neuroscience* 1: 5, 2007.

Bregman AS. *Auditory Scene Analysis: The Perceptual Organization of Sound*. The MIT Press, Cambridge, 1990.

Brouwer S, Van Engen KJ, Calandruccio L, and Bradlow AR. Linguistic contributions to speech-on-speech masking for native and non-native listeners: language familiarity and semantic content. *The Journal of the Acoustical Society of America* 131: 1449-1464, 2012.

Calandruccio L, Dhar S, and Bradlow AR. Speech-on-speech masking with variable access to the linguistic content of the masker speech. *The Journal of the Acoustical Society of America* 128: 860-869, 2010.

Campbell T, Kerlin JR, Bishop CW, and Miller LM. Methods to eliminate stimulus transduction artifact from insert earphones during electroencephalography. *Ear and hearing* 33: 144-150, 2012.
Carabellese C, Appollonio I, Rozzini R, Bianchetti A, Frisoni GB, Frattola L, and Trabucchi M. Sensory impairment and quality of life in a community elderly population. *Journal of the American Geriatrics Society* 41: 401-407, 1993.
Caspary DM, Hughes LF, Schatteman TA, and Turner JG. Age-related changes in the response properties of cartwheel cells in rat dorsal cochlear nucleus. *Hearing research* 216-217: 207-215, 2006.
Caspary DM, Ling L, Turner JG, and Hughes LF. Inhibitory neurotransmission, plasticity and aging in the mammalian central auditory system. *The Journal of experimental biology* 211: 1781-1791, 2008.
Caspary DM, Milbrandt JC, and Helfert RH. Central auditory aging: GABA changes in the inferior colliculus. *Experimental gerontology* 30: 349-360, 1995.
Caspary DM, Schatteman TA, and Hughes LF. Age-related changes in the inhibitory response properties of dorsal cochlear nucleus output neurons: role of inhibitory inputs. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 25: 10952-10959, 2005.
Chambers Anna R, Resnik J, Yuan Y, Whitton Jonathon P, Edge Albert S, Liberman MC, and Polley Daniel B. Central Gain Restores Auditory Processing following Near-Complete Cochlear Denervation. *Neuron* 89: 867-879, 2016.
Chandrasekaran B, and Kraus N. The scalp-recorded brainstem response to speech: neural origins and plasticity. *Psychophysiology* 47: 236-246, 2010.
Cheyne D, and Papanicolaou A. *Magnetoencephalography and Magnetic Source Imaging.* Oxford University Press, 2006.
Chimento T, and Schreiner C. Selectively eliminating cochlear microphonic contamination from the frequency-following response. *Electroencephalogr Clin Neurophysiol* 75: 88-96, 1990.
Choi I, Rajaram S, Varghese LA, and Shinn-Cunningham BG. Quantifying attentional modulation of auditory-evoked cortical responses from single-trial electroencephalography. *Frontiers in human neuroscience* 7: 115, 2013.
Clinard CG, and Cotter CM. Neural representation of dynamic frequency is degraded in older adults. *Hearing research* 323: 91-98, 2015.
Clinard CG, and Tremblay KL. Aging degrades the neural encoding of simple and complex sounds in the human brainstem. *Journal of the American Academy of Audiology* 24: 590-599; quiz 643-594, 2013.
Clinard CG, Tremblay KL, and Krishnan AR. Aging alters the perception and physiological representation of frequency: evidence from human frequency-following response recordings. *Hearing research* 264: 48-55, 2010.
Collier R, and Hart Jt. The Role of Intonation in Speech Perception. In: *Structure and Process in Speech Perception: Proceedings of the Symposium on Dynamic Aspects of Speech Perception held at IPO, Eindhoven, Netherlands, August 4–6, 1975*, edited by Cohen A, and Nooteboom SG. Berlin, Heidelberg: Springer Berlin Heidelberg, 1975, p. 107-123.
David SV, Mesgarani N, Fritz JB, and Shamma SA. Rapid synaptic depression explains nonlinear modulation of spectro-temporal tuning in primary auditory cortex by

natural stimuli. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 29: 3374-3386, 2009.
David SV, Mesgarani N, and Shamma SA. Estimating sparse spectro-temporal receptive fields with natural stimuli. *Network (Bristol, England)* 18: 191-212, 2007.
de Cheveigne A, and Simon JZ. Denoising based on spatial filtering. *J Neurosci Methods* 171: 331-339, 2008.
De Cheveigné A, and Simon JZ. Denoising based on time-shift PCA. *J Neurosci Methods* 165: 297-305, 2007.
de Villers-Sidani E, Alzghoul L, Zhou X, Simpson KL, Lin RC, and Merzenich MM. Recovery of functional and structural age-related changes in the rat primary auditory cortex with operant training. *Proceedings of the National Academy of Sciences of the United States of America* 107: 13900-13905, 2010.
Delorme A, and Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of neuroscience methods* 134: 9-21, 2004.
Ding N, Chatterjee M, and Simon JZ. Robust cortical entrainment to the speech envelope relies on the spectro-temporal fine structure. *NeuroImage* 88c: 41-46, 2013.
Ding N, and Simon JZ. Adaptive temporal encoding leads to a background-insensitive cortical representation of speech. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 33: 5728-5735, 2013.
Ding N, and Simon JZ. Emergence of neural encoding of auditory objects while listening to competing speakers. *Proceedings of the National Academy of Sciences of the United States of America* 109: 11854-11859, 2012.
Don M, and Eggermont JJ. Analysis of the click-evoked brainstem potentials in man using high-pass noise masking. *The Journal of the Acoustical Society of America* 63: 1084-1092, 1978.
Fitzgibbons PJ, and Gordon-Salant S. Aging and temporal discrimination in auditory sequences. *The Journal of the Acoustical Society of America* 109: 2955-2963, 2001.
Fitzgibbons PJ, and Gordon-Salant S. Auditory temporal processing in elderly listeners. *Journal of the American Academy of Audiology* 7: 183-189, 1996.
Frisina DR, and Frisina RD. Speech recognition in noise and presbycusis: relations to possible neural mechanisms. *Hearing research* 106: 95-104, 1997.
Fritz J, Shamma S, Elhilali M, and Klein D. Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. *Nature neuroscience* 6: 1216-1223, 2003.
Furman AC, Kujawa SG, and Liberman MC. Noise-induced cochlear neuropathy is selective for fibers with low spontaneous rates. *Journal of neurophysiology* 110: 577-586, 2013.
Galambos R, Makeig S, and Talmachoff PJ. A 40-Hz auditory potential recorded from the human scalp. *Proceedings of the National Academy of Sciences of the United States of America* 78: 2643-2647, 1981.
Galbraith GC, Threadgill MR, Hemsley J, Salour K, Songdej N, Ton J, and Cheung L. Putative measure of peripheral and brainstem frequency-following in humans. *Neuroscience letters* 292: 123-127, 2000.

Gates GA, Cobb JL, Linn RT, Rees T, Wolf PA, and D'Agostino RB. Central auditory dysfunction, cognitive dysfunction, and dementia in older people. *Archives of otolaryngology--head & neck surgery* 122: 161-167, 1996.
Getzmann S, Golob EJ, and Wascher E. Focused and divided attention in a simulated cocktail-party situation: ERP evidence from younger and older adults. *Neurobiology of aging* 2016.
Getzmann S, and Naatanen R. The mismatch negativity as a measure of auditory stream segregation in a simulated "cocktail-party" scenario: effect of age. *Neurobiology of aging* 36: 3029-3037, 2015.
Gordon-Salant S, and Fitzgibbons PJ. Temporal factors and speech recognition performance in young and elderly listeners. *Journal of speech and hearing research* 36: 1276-1285, 1993.
Gordon-Salant S, Yeni-Komshian GH, Fitzgibbons PJ, and Barrett J. Age-related differences in identification and discrimination of temporal cues in speech segments. *The Journal of the Acoustical Society of America* 119: 2455-2466, 2006.
Gorga M, Abbas P, and Worthington D. Stimulus calibration in ABR measurements. In: The auditory brainstem response. *(Jacobsen J, ed)* 49-62. San Diego: College Hill, 1985.
Greenberg S, Marsh JT, Brown WS, and Smith JC. Neural temporal coding of low pitch. I. Human frequency-following responses to complex tones. *Hear Res* 25: 91-114, 1987.
Greenberg S, and Takayuki A. What are the essential cues for understanding spoken language? *IEICE transactions on information and systems* 87: 1059-1070, 2004.
He NJ, Mills JH, Ahlstrom JB, and Dubno JR. Age-related differences in the temporal modulation transfer function with pure-tone carriers. *The Journal of the Acoustical Society of America* 124: 3841-3849, 2008.
Henry KS, Kale S, and Heinz MG. Noise-induced hearing loss increases the temporal precision of complex envelope coding by auditory-nerve fibers. *Frontiers in Systems Neuroscience* 8: 20, 2014.
Herbst KG, and Humphrey C. Hearing impairment and mental state in the elderly living at home. *British medical journal* 281: 903-905, 1980.
Hughes LF, Turner JG, Parrish JL, and Caspary DM. Processing of broadband stimuli across A1 layers in young and aged rats. *Hearing research* 264: 79-85, 2010.
Humes LE, and Christopherson L. Speech identification difficulties of hearing-impaired elderly persons: the contributions of auditory processing deficits. *Journal of speech and hearing research* 34: 686-693, 1991.
Humes LE, Garner CB, Wilson DL, and Barlow NN. Hearing-aid outcome measured following one month of hearing aid use by the elderly. *Journal of speech, language, and hearing research : JSLHR* 44: 469-486, 2001.
Humes LE, and Roberts L. Speech-recognition difficulties of the hearing-impaired elderly: the contributions of audibility. *Journal of speech and hearing research* 33: 726-735, 1990.
Juarez-Salinas DL, Engle JR, Navarro XO, and Recanzone GH. Hierarchical and serial processing in the spatial auditory cortical pathway is degraded by natural aging. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 30: 14795-14804, 2010.

Kay DWK, Beamish P, and Roth M. Old Age Mental Disorders in Newcastle upon Tyne: Part I: A Study of Prevalence. In: *The British Journal of Psychiatry*1964, p. 146-158.
Killion MC, Niquette PA, Gudmundsen GI, Revit LJ, and Banerjee S. Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America* 116: 2395-2405, 2004.
King C, Warrier CM, Hayes E, and Kraus N. Deficits in auditory brainstem pathway encoding of speech sounds in children with learning problems. *Neurosci Lett* 319: 111-115, 2002.
Klatt DH. Software for a cascade/parallel formant synthesizer *The Journal of the Acoustical Society of America* 67: 971-995, 1980.
Lachaux JP, Lutz A, Rudrauf D, Cosmelli D, Le Van Quyen M, Martinerie J, and Varela F. Estimating the time-course of coherence between single-trial brain signals: an introduction to wavelet coherence. *Neurophysiologie clinique = Clinical neurophysiology* 32: 157-174, 2002.
Laforge RG, Spector WD, and Sternberg J. The relationship of vision and hearing impairment to one-year mortality and functional decline. *Journal of Aging Health* 4: 126-148, 1992.
Larsby B, Hallgren M, and Lyxell B. The interference of different background noises on speech processing in elderly hearing impaired subjects. *International journal of audiology* 47 Suppl 2: S83-90, 2008.
Lash A, Rogers CS, Zoller A, and Wingfield A. Expectation and entropy in spoken word recognition: effects of age and hearing acuity. *Experimental aging research* 39: 235-253, 2013.
Lee CC, and Middlebrooks JC. Auditory cortex spatial sensitivity sharpens during task performance. *Nature neuroscience* 14: 108-114, 2011.
Liden G, and Korsan-Bengtsen M. Audiometric manifestations of retrocochlear lesions. *Adv Otorhinolaryngol* 20: 271-287, 1973.
Lin FR, Thorpe R, Gordon-Salant S, and Ferrucci L. Hearing loss prevalence and risk factors among older adults in the United States. *The journals of gerontology Series A, Biological sciences and medical sciences* 66: 582-590, 2011.
Lin FR, Yaffe K, Xia J, Xue QL, Harris TB, Purchase-Helzner E, Satterfield S, Ayonayon HN, Ferrucci L, and Simonsick EM. Hearing loss and cognitive decline in older adults. *JAMA internal medicine* 173: 293-299, 2013.
Lister JJ, Maxfield ND, Pitt GJ, and Gonzalez VB. Auditory evoked response to gaps in noise: older adults. *International journal of audiology* 50: 211-225, 2011.
McDermott JH. The cocktail party problem. *Current Biology* 19: R1024-R1027, 2009.
Michelini S, Arslan E, Prosser S, and Pedrielli F. Logarithmic display of auditory evoked potentials. *Journal of biomedical engineering* 4: 62-64, 1982.
Nasreddine ZS, Phillips NA, Bedirian V, Charbonneau S, Whitehead V, Collin I, Cummings JL, and Chertkow H. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society* 53: 695-699, 2005.
Olsen WO, Noffsinger D, and Kurdziel S. Acoustic reflex and reflex decay. Occurrence in patients with cochlear and eighth nerve lesions. *Arch Otolaryngol* 101: 622-625, 1975.

Otto WC, and McCandless GA. Aging and the auditory brain stem response. *International journal of audiology* 21: 466-473, 1982.
Overton JA, and Recanzone GH. Effects of aging on the response of single neurons to amplitude modulated noise in primary auditory cortex of Rhesus macaque. *Journal of neurophysiology* jn.01098.02015, 2016.
Parbery-Clark A, Anderson S, Hittner E, and Kraus N. Musical experience offsets age-related delays in neural timing. *Neurobiology of aging* 33: 1483.e1481-1484, 2012.
Parbery-Clark A, Marmel F, Bair J, and Kraus N. What subcortical-cortical relationships tell us about processing speech in noise. *The European journal of neuroscience* 33: 549-557, 2011.
Parthasarathy A, and Bartlett EL. Age-related auditory deficits in temporal processing in F-344 rats. *Neuroscience* 192: 619-630, 2011.
Parthasarathy A, Cunningham PA, and Bartlett EL. Age-related differences in auditory processing as assessed by amplitude-modulation following responses in quiet and in noise. *Frontiers in aging neuroscience* 2: 152, 2010.
Peelle JE, Troiani V, Grossman M, and Wingfield A. Hearing loss in older adults affects neural systems supporting speech comprehension. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 31: 12638-12643, 2011.
Peelle JE, Troiani V, Wingfield A, and Grossman M. Neural processing during older adults' comprehension of spoken sentences: age differences in resource allocation and connectivity. *Cereb Cortex* 20: 773-782, 2010.
Pichora-Fuller MK, and Schneider BA. Masking-level differences in the elderly: a comparison of antiphasic and time-delay dichotic conditions. *Journal of speech and hearing research* 34: 1410-1422, 1991.
Pichora-Fuller MK, Schneider BA, and Daneman M. How young and old adults listen to and remember speech in noise. *The Journal of the Acoustical Society of America* 97: 593-608, 1995.
Pichora-Fuller MK, Schneider BA, Macdonald E, Pass HE, and Brown S. Temporal jitter disrupts speech intelligibility: a simulation of auditory aging. *Hearing research* 223: 114-121, 2007.
Presacco A, Jenkins K, Lieberman R, and Anderson S. Effects of Aging on the Encoding of Dynamic and Static Components of Speech. *Ear and hearing* 36: e352-363, 2015.
Presacco A, Simon JZ, and Anderson S. Evidence of degraded representations of speech in noise in the aging midbrain and cortex. *Journal of neurophysiology Under review*: 2016 "co-submitted to J. Neurophysiol. along with this manuscript".
Recanzone GH, Engle JR, and Juarez-Salinas DL. Spatial and temporal processing of single auditory cortical neurons and populations of neurons in the macaque monkey. *Hearing research* 271: 115-122, 2011.
Rogers CS, Jacoby LL, and Sommers MS. Frequent false hearing by older adults: the role of age differences in metacognition. *Psychology and aging* 27: 33-45, 2012.
Rogers CS, and Wingfield A. Stimulus-independent semantic bias misdirects word recognition in older adults. *The Journal of the Acoustical Society of America* 138: El26-30, 2015.
Ross B, Schneider B, Snyder JS, and Alain C. Biological Markers of Auditory Gap Detection in Young, Middle-Aged, and Older Adults. *PLoS ONE* 5: e10101, 2010.

Särelä J, and Valpola H. Denoising Source Separation. *J Mach Learn Res* 6: 233-272, 2005.
Schatteman TA, Hughes LF, and Caspary DM. Aged-related loss of temporal processing: altered responses to amplitude modulated tones in rat dorsal cochlear nucleus. *Neuroscience* 154: 329-337, 2008.
Schneider BA, and Hamstra SJ. Gap detection thresholds as a function of tonal duration for younger and older listeners. *The Journal of the Acoustical Society of America* 106: 371-380, 1999.
Shamma SA, Elhilali M, and Micheyl C. Temporal coherence and attention in auditory scene analysis. *Trends in neurosciences* 34: 114-123, 2011.
Slee SJ, and David SV. Rapid Task-Related Plasticity of Spectrotemporal Receptive Fields in the Auditory Midbrain. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 35: 13090-13102, 2015.
Smith JC, Marsh JT, and Brown WS. Far-field recorded frequency-following responses: evidence for the locus of brainstem sources. *Electroencephalography and clinical neurophysiology* 39: 465-472, 1975.
Soros P, Teismann IK, Manemann E, and Lutkenhoner B. Auditory temporal processing in healthy aging: a magnetoencephalographic study. *BMC Neurosci* 10: 34, 2009.
Suga N. Role of corticofugal feedback in hearing. *Journal of comparative physiology A, Neuroethology, sensory, neural, and behavioral physiology* 194: 169-183, 2008.
Surprenant AM. Effects of noise on identification and serial recall of nonsense syllables in older and younger adults. *Neuropsychology, development, and cognition Section B, Aging, neuropsychology and cognition* 14: 126-143, 2007.
Tallon-Baudry C, Bertrand O, Delpuech C, and Permier J. Oscillatory gamma-band (30-70 Hz) activity induced by a visual search task in humans. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 17: 722-734, 1997.
Tang X, Zhu X, Ding B, Walton J, Frisina R, and Su J. Age-related hearing loss: GABA, nicotinic acetylcholine and NMDA receptor expression changes in spiral ganglion neurons of the mouse. *Neuroscience* 259: 184-193, 2014.
Tremblay KL, Piskosz M, and Souza P. Effects of age and age-related hearing loss on the neural representation of speech cues. *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology* 114: 1332-1343, 2003.
Tun PA, McCoy S, and Wingfield A. Aging, hearing acuity, and the attentional costs of effortful listening. *Psychology and aging* 24: 761-766, 2009.
Tun PA, O'Kane G, and Wingfield A. Distraction by competing speech in young and older adult listeners. *Psychology and aging* 17: 453-467, 2002.
Uhlmann RF, Larson EB, Rees TS, Koepsell TD, and Duckert LG. Relationship of hearing impairment to dementia and cognitive dysfunction in older adults. *JAMA* 261: 1916-1919, 1989.
Vander Werff KR, and Burns KS. Brain stem responses to speech in younger and older adults. *Ear and hearing* 32: 168-180, 2011.
Varghese L, Bharadwaj HM, and Shinn-Cunningham BG. Evidence against attentional state modulating scalp-recorded auditory brainstem steady-state responses. *Brain research* 1626: 146-164, 2015.

Waldert S, Pistohl T, Braun C, Ball T, Aertsen A, and Mehring C. A review on directional information in neural signals for brain-machine interfaces. *Journal of physiology, Paris* 103: 244-254, 2009.
Wallace MN, Rutkowski RG, Shackleton TM, and Palmer AR. Phase-locked responses to pure tones in guinea pig auditory cortex. *Neuroreport* 11: 3989-3993, 2000.
Walton JP, Frisina RD, Ison JR, and O'Neill WE. Neural correlates of behavioral gap detection in the inferior colliculus of the young CBA mouse. *Journal of comparative physiology A, Sensory, neural, and behavioral physiology* 181: 161-176, 1997.
Walton JP, Frisina RD, and O'Neill WE. Age-related alteration in processing of temporal sound features in the auditory midbrain of the CBA mouse. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 18: 2764-2776, 1998.
Wang H, Turner JG, Ling L, Parrish JL, Hughes LF, and Caspary DM. Age-related changes in glycine receptor subunit composition and binding in dorsal cochlear nucleus. *Neuroscience* 160: 227-239, 2009.
Ward CM, Rogers CS, Van Engen KJ, and Peelle JE. Effects of Age, Acoustic Challenge, and Verbal Working Memory on Recall of Narrative Speech. *Experimental aging research* 42: 126-144, 2016.
Willott JF. Anatomic and physiologic aging: a behavioral neuroscience perspective. *Journal of the American Academy of Audiology* 7: 141-151, 1996.
Wronka E, Kaiser J, and Coenen AM. Neural generators of the auditory evoked potential components P3a and P3b. *Acta neurobiologiae experimentalis* 72: 51-64, 2012.
Zhu J, and Garcia E. The Wechsler Abbreviated Scale of Intelligence (WASI). *New York: Psychological Corporation* 1999.